GHENT UNIVERSITY

MASTER THESIS

Opinion dynamics on social media with stubborn actors

Author:
Nina BOTTE

Supervisor:
Prof. Dr. Luis E.C. ROCHA

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Physics and Astronomy

in the

Faculty of SCIENCE
Department of PHYSICS AND ASTRONOMY

Academic year 2020-2021

Acknowledgments

Abstract

The past decade, the rise of social media has become undeniable. More and more people are getting engaged on some type of social media platform and hence, it is probably not wrong to say that nowadays most of us have, beside our real life, a virtual life on social media. This rapidly increased engagement on social media undoubtedly has many advantages. However, every upside has a downside and social media is no different. This thesis deals with the impact of social media on the formation and evolution of opinions in society.

Social media allows for the spread of information at a faster pace and at an unprecedented scale. However, people are still bounded by time- and cognitive constraints. In order to handle the information overflow, social media companies design filtering algorithms that can fine-tune and individualize the information someone is exposed to. This filtering and individualizing of information may have some potentially dangerous downfalls in the sense that people might get trapped in "echo chambers" that confirm their own beliefs. It is important to know the effects of such algorithmic personalization on the appearance and formation of phenomena such as polarization, spread of extremism, formation of echo chambers, etc.

People are, however, more than sheep that thoughtlessly follow their shepherd (contrary to what some books might lead us to believe). They are capable of critical thinking and may be hesitant to change their current opinion. This aspect of potential stubbornness should not be ignored.

In this thesis toy models are designed to investigate the interplay between network structure, filtering algorithms and individual resistance to change in the evolution of two competing opinions.

We will investigate whether network properties such as community structure and clustering coefficient may enhance or hamper the formation of echo chambers. We will use some basic filtering algorithms to determine their possible effects on the appearance of echo chambers and we will investigate the impact of individual stubbornness.

Beside that, we will also pay some attention to the possible ways of distributing the opinions and the stubborn people and how this may effect the opinion dynamics in the system.

Finally, we will test the model on real-world networks. Some question that may arise are: Do the results agree with those obtained from the network models? What are the similarities or differences in the results? What are the possible underlying reason/structures for these similarities/differences?

Samenvatting

Table of Contents

List of Figures

Chapter 1

Introduction

The past decade social media has become ever more important in our daily lives. Nowadays most of us have, beside a real-world life, a virtual life on social media. This increased engagement on social media may have a huge impact on the formation of opinions, not only on the individual scale, but also on a broader, societal scale.

It is known that individuals not only form their opinions through self-reflection, but also through interactions with other people and with their surroundings [45]. The broad range of interactions people undergo on social media with people from all over the world can thus not be underestimated in the process of opinion formation. The understanding of the role of social media on the emergence of polarization and extremism in society is of uttermost importance.

It is also known that some people change their opinions more easily than others, that is, once in a while you encounter somebody that is really stubborn and persistent towards its own opinion. One of the aims of this thesis is to design toy models of opinion dynamics with stubborn actors on theoretical and real-world social networks to analyze the interplay between individual resistance to change and the network structure in the evolution of two competing opinions.

One important feature of opinion dynamics on social media is the use of algorithmic personalization. Social media allow for simultaneous or asynchronous interactions between people without geographical constraints and thus allow for the spread of information at a faster pace and at an unprecedented scale [45]. People are, however, still bound by temporal and cognitive constraints. In order to assure a more pleasant/convenient experience on their platforms, social media companies use algorithms to order/filter the posts on an individual's time-line according to what might be relevant to that individual [45]. It is of great importance to know the possible effects of these filtering algorithms on the evolution of opinions in the population.

Opinion dynamics models have two important layers. On one hand, we introduce social networks to describe the underlying structure of social interactions; on the other hand, we need appropriate opinion dynamics/formation models to reproduce the opinion dynamics and opinion formation processes found in real life.

Section 1.1 gives an introduction to the first layer, namely that of complex networks. The second layer, the opinion dynamics layer, is introduced in Section 1.2. Section 1.3 introduces concepts such as statistical and social physics. Finally, in Section 1.4, the hypotheses and goals of this thesis are discussed.

1.1   Complex networks

Complex networks have become more and more popular for analyzing complex, dynamical systems. They are used in fields such as physics, economics, social sciences, biology, etc [15]. These different fields are very diverge, but they have at least one common ground: they often deal with a large number of variables/components that interact with each other. In other words, in all these fields one encounters complex systems, systems where it is not possible to predict collective behavior based on the properties of the individual components alone [17]. Complex systems often display phenomena such as non-linearity, emergence, spontaneous order,...

One of the tools to deal with these complex systems and to give us more insight in the possible underlying structures are complex networks. A network, also often called a graph (both terms will be used in this thesis), is a structure composed of nodes or vertices and a set of links or edges that indicate the interactions between the nodes (in this thesis terms like agent, actor, user, individual, people, etc will be interchangebly used to denote the nodes; edges will usually just be called edges) [17].

Representing/modeling a complex system as a graph makes the system appear more simple and tractable, while it still includes the non-linearity of these systems. One can find the language to describe networks in mathematical graph theory. However, complex systems in real-life situations often deal with a huge number of components, so the use of statistical and high-performance computing tools is inevitable [17]. One could argue that the study of complex networks lies somewhere at the intersection of mathematical graph theory and statistical physics [16].

The advantage of working with network models is that they reduce the level of complexity encountered in the real world, so that one can treat these systems in a more practical way. However, we still want our models to display properties similar to the ones seen in real systems [17]. Since many real systems are not static but evolve in time, this means that we don't only need to deal with static networks but also with temporal networks. Temporal networks are networks where the edges are not always active, but instead become active for some periods of time [30]. In temporal networks time becomes an explicit element in the network representation [30]. Static networks have been widely studied and are often convenient for their analytic tractability, whereas temporal networks are, in some cases, more realistic [17].

Some other important concepts in network theory are the degree distribution, clustering/community structure, connectivity, etc. These concepts will be explained in depth in Chapter 2, Section 2.1, but let us already anticipate a bit on the case of social networks. It is found that many real-life social systems have a power-law degree distribution [39]. This heterogeneous or scale-free degree distribution represents one of the three general properties of social networks. The other two are short distances and high clustering. These two are also referred to as "small world phenomenon" (note that in the literature the term "small world" sometimes only refers to the property of short distances, however, in this thesis, we will use it to indicate both a high clustering and short distances) [39]. Ideally, our theoretical network

should exhibit these three properties. Some theoretical networks, that are thoroughly studied, do not possess all of these three properties. However, they might still be used, because of their simplicity and ability to produce analytic results. When using these models one must always bear in mind their limitations to reproduce some properties encountered in real social systems.

One of the goals of this thesis is to investigate/determine the role of the underlying network structure on the formation and dynamics of opinions in the system. Some examples of theoretical network models used in this thesis are the Erdős-Rényi network, the stochastic block model, the Watts-Strogatz model, etc. These will be explained thoroughly in Chapter 2, Section 2.2.

1.2   Opinion dynamics

As stated before, the formation of an individual's opinion is the result of the interplay of many factors such as self-reflection, peer pressure, the individual's personality (eg. stubbornness), the information someone is exposed to, etc. Opinion dynamics models should try to capture these complex processes in a simplified way. Several models have been developed ranging from simple binary models to more complex, continuous approaches [47].
The basic idea of all opinion dynamics models is that the nodes or agents in a social network have a variable that represents their opinion and that is updated according to some predefined rules. These models are obviously a simplification of real-world opinion dynamics. It is however shown that they do display a lot of aspects of real opinion formation such as agreement, transitions between order (consensus) and disorder (fragmentation), polarization, formation of echo chambers (clusters of people with the same opinion), etc [47].

The opinion of the agents in the model can be either discrete or continuous. This thesis deals with models where each agent can have one of the two opinions A or B and thus only deals with the case of discrete opinions. This might come across as a huge simplification of the real-world complexity of opinion formation, but also in real life people often have to choose between two competing opinions (eg. republicans or democrats, cat or dog, renting or buying a house,...).

The rules that determine how an agent updates his or her opinion can include many different aspects. They can, for example, be as simple as a majority rule (i.e. the majority model: if the majority of your neighbors have a certain opinion, you adopt that opinion as well) or can include a more probabilistic way of updating. A small overview of possible opinion dynamics models is given in Chapter 3, Section 3.1 and Section 3.2.
One can also include a resistance parameter that determines the hesitation of an agent to change to a new opinion and/or a parameter that determines the influence of an agent on others. Furthermore, the agents can have a parameter that determines whether they are active or not. The introduction of stubborn people will be discussed in Chapter 3, Section 3.5; the activation mechanism of nodes is discussed in Chapter 3, Section 3.4.

Since this thesis deals with the particular case of opinion dynamics on on-line social platforms, it is important to define/set up different filtering algorithms. The filtering algorithms that real social media

companies use, are corporate secrets, but we know that there are three main principles of content curation: popularity, semantic and collaborative filtering [45]. Popularity filtering refers to the practice of promoting content that is popular across the platform; semantic filtering means that post similar to previous consumed posts are recommended and collaborative filtering suggests posts that are similar to the ones our friends consume [45]. In Chapter 3, Section 3.3 a detailed description of the filtering algorithms used in this thesis will be given.

1.3   Statistical and social physics

Another important tool in the study of opinion dynamics/formations in large groups of people is the use of statistical physics. This branch of physics offers a framework that relates microscopic properties of atoms, molecules,... to macroscopic observed behavior and has a wide field of applications inside and outside the world of physics.

The observation that large number of people display collective, 'macroscopic' behavior begged for the use of the concepts and insights developed in statistical physics [47]. The application of the theory of statistical physics on social phenomena is referred to as social physics or sociophysics. The 'microscopic' constituents are now individual humans who interact with a limited number of other individuals and in that way form complex, 'macroscopic' groups such as human societies. These 'macroscopic' groups display stunning regularities, transitions from disorder to order, the emergence of consensus, universality, etc [9]. The statistical physics approach of these social systems tries to explain the found regularities at large scale as collective effects of the interactions among these individuals [47].

The statistical approach of social phenomena has one big difficulty compared to the description of gases and other physical phenomena. In regular physics, the microscopic constituents are 'simple' atoms or molecules, whose dynamics and interactions with each other are often well known and can be put in relatively simple laws with only a few parameters [10]. Humans, on the other hand, are far from simple entities and the dynamics of a single individual are far from known [10]. Even if the dynamics/behavior of individuals would be well know, then the interactions between them would still be complicated and it would be impossible to describe them with simple laws and only a few parameters [10]. Hence, it may seem that extracting useful and qualitative information out of social models is a nearly impossible and almost hopeless task [10].

However, the use of statistical physics may also offer a way out. In most situations the qualitative (and sometimes even quantitative) properties of large scale phenomena do not depend on the microscopic details of the process [10]. Instead, the relevant aspects for the global behavior are often higher level features such as symmetries, conservation laws, dimensionality, etc [10]. This introduces a concept of universality. With this universality in mind, one can proceed to build models of human/social behavior where only the simplest and most important properties of single individuals and their interactions are taken into account [10]. An important factor herein is the comparison with real/empirical data to investigate whether or not the trends seen in real life are compatible with the outputs generated by the models [10]. Note that this interplay between theory, model, experiment and real-life data is not only an important factor in social physics, but also in regular physical research and science in general.

1.4   This thesis: hypotheses and goals

As said before, this thesis will try to investigate the interplay between the social network structure and the individual resistance to change in the evolution of two competing opinions. On one hand, we want to determine the influence of different network structures on the formation and evolution of opinions, whereas, on the other hand, we would like to investigate the effect of stubborn agents in the network. Since this thesis deals with opinion dynamics on social media, the effect of filtering algorithms must also be included.

In one of their previous papers (*"Modelling opinion dynamics in the age of algorithmic personalization"*, see reference [45]) prof. Nicola Perra and prof. Luis E.C. Rocha investigated both the effect of different filtering algorithms and different network topologies on the formation and evolution of two competing opinions. They also considered the effect of nudging (meaning that one opinion is pushed to all agents in the network).

Their main findings were that the algorithmic filtering could not break the status quo when the prevalence of opinions was equally distributed in the population and that topological correlations (such as high clustering coefficient) could result in the formation of echo chambers [45]. On the other hand, a more heterogeneous contact pattern could hamper this formation of echo chambers [45]. If the two opinions were, initially, not equally distributed, one of the filtering algorithms (namely semantic filtering) caused a further increase in the predominant opinion [45]. In the case of nudging, they found that the opinions in the population moved to the nudged opinion relatively fast, even in the case of small nudging [45].

They, however, did not investigate the effect of stubborn agents in the network, nor did they investigate networks with a community structure. This thesis will build on their model and incorporate these new features. We will compare networks that have a community structure and relatively low clustering to networks without community structure, but with a high clustering coefficient (a detailed description of concepts like clustering coefficient and community structure will be given in Chapter 2, Section 2.1). We want to investigate whether networks with community structure are able to form echo chambers, such as one can observe in networks with a high clustering coefficient and, if so, if one can, for example, observe convergence to one opinion inside the communities. We will also pay some attention to networks that have both a high clustering coefficient and a community structure. Can we observe echo chambers in those kinds of networks? If yes, can we determine which property might be the main driver for the formation of echo chambers: high clustering or community structure?

On the other hand, we want to determine the impact of stubborn agents on the formation and evolution of opinions in the network. Some interesting questions that may come to mind are: Is there an interplay between network structure and stubbornness? What about the possible interplay between the filtering algorithms and the introduction of stubborn people? Does the distribution of stubborn people play a role (eg. randomly distributed versus clustered together inside communities)?

Chapter 2

Networks

2.1 Definition and network measurements

Networks or graphs are conceptually very simple and flexible objects that are made of a set of nodes/vertices V and a set of edges/links E, where the elements of E determine connections between elements of V [14]

$$E \subseteq \{\{x,y\}|x,y \in V\} \, . \tag{2.1}$$

If self-loops are not allowed, we need the following condition on the elements of E

$$E \subseteq \{\{x,y\}|x,y \in V \text{ and } x \neq y\} \, . \tag{2.2}$$

A graph with no self-loops is a simple graph. The nodes x and y of an edge $\{x,y\}$ are called the endpoints of the edge. It is possible that nodes are not joined by any edge, such nodes are called isolated or disconnected. If two or more edges have the same pair of endpoints, the graph is called a multi-graph. This thesis is not concerned with multi-graphs, nor are self-loops allowed. Graphs in which the edges have an orientation are called directed graphs. If the edges have weights w_{ij}, one speaks of a weighted graph.
A graph can be represented in different ways. One of the most common ways is by use of the adjacency matrix.

2.1.1 Adjacency matrix

The adjacency matrix $\mathbf{A}$ is a $N \times N$ matrix (for a graph with N nodes) where the elements indicate whether two nodes are connected by an edge or not [17]

$$A_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

For a simple, undirected graph, the adjacency matrix is symmetric ($A_{ij} = A_{ji}$) with zeros on the diagonal ($A_{ii} = 0$). Directed graphs can have asymmetric adjacency matrices. If we are dealing with

weighted graphs, the elements of the adjacency matrix are the weights w_{ij} of the edges

$$A_{ij} = \begin{cases} w_{ij} & \text{if nodes } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise,} \end{cases} \tag{2.4}$$

with, generally, $0 \leqslant w_{ij} \leqslant 1$ [17].

2.1.2 Degree and degree distribution

The adjacency matrix contains a lot of information such as the degree of a node. The degree k_i of a node i is the number of edges attached to node i or, in others words, k_i represents the number of nearest neighbors of node i. The degree can be easily obtained from the adjacency matrix [17]

$$k_i = \sum_{j=1}^{N} A_{ij} \ . \tag{2.5}$$

In the case of a directed graph one can differentiate between the number of incoming edges k_i^{in} and the number of outgoing edges k_i^{out} of node i. These are called the in-degree and out-degree respectively and are defined as [17]

$$k_i^{\text{in}} = \sum_{j=1}^{N} A_{ji} \ , \qquad\qquad k_i^{\text{out}} = \sum_{j=1}^{N} A_{ij} \ . \tag{2.6}$$

The total degree of node i is then $k_i = k_i^{\text{in}} + k_i^{\text{out}}$. For weighted graphs the degree is easily generalized to the weighted degree s_i, often called strength [32]

$$s_i = \sum_{j=1}^{N} w_{ij} \ . \tag{2.7}$$

From the degree of the nodes in the network one can construct the degree distribution. The degree distribution $P(k)$ represents the fraction of nodes with a degree k or, in other words, the degree distribution gives the probability that a randomly selected node has a degree k. The average degree can be obtained from the degree distribution [17]

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i = \sum_k k P(k) \ , \tag{2.8}$$

where N is the total number of nodes in the network. The degree distribution also allows us to classify networks. The two most important classes are homogeneous and heterogeneous networks. Homogeneous networks have a bell curved degree distribution, eg. a Poisson distribution. In this case, most of the nodes have a degree close to the average degree $\langle k \rangle$. Heterogeneous networks, on the other hand, have a power-law degree distribution, $P(k) \sim k^{-\gamma}$. They are also referred to as scale-free networks, since they do not posses a characteristic length scale or, in this case, the average degree isn't a characteristic scale for the network [17].

2.1.3 Connectivity

In physics literature, the degree of a node is often called the connectivity [16][19]. This is different from the mathematical/graph definition of connectivity, where connectivity refers to the minimum number of nodes or edges that need to be removed to separate the remaining nodes in isolated subgraphs.

Two nodes u and v of a graph are connected if there exist a path between them. If the two nodes are connected by a path that contains only one edge, the two nodes are called adjacent.

The average degree $\langle k \rangle$ is a global measurement of the connectivity of the network [15]. If the average connectivity $\langle k \rangle$ is too small (this is, if there are too few edges), there will be many isolated nodes and only a few clusters with a small number of nodes. If more and more edges are added to the network, the small clusters will grow and will tend to connect to each other to form larger clusters. At some critical value of the connectivity, most nodes will be connected into a giant cluster (the giant component), which characterizes the percolation of the network [16].

2.1.4 Assortative and disassortative networks

Another concept that is closely related to the degree of a node is the notion of assortative and disassortative networks. More precisely, the notion of assortativity and disassortativity has to do with degree correlations.

The joint degree distribution $P(k, k')$ is the probability that an arbitrary edge connects a node with degree k to a node with degree k' [16]. The conditional probability $P(k|k')$, which is another way to express dependencies between node degrees, is the probability that an arbitrary neighbor of a node with degree k has a degree k' [16]

$$P(k|k') = \frac{\langle k \rangle P(k, k')}{k P(k)} \ .$$

(2.9)

The average degree of the nearest neighbors of nodes with degree k can be computed as [16]

$$k_{nn}(k) = \sum_{k'} k' P(k'|k) \ .$$

(2.10)

If there are no degree correlations, k_{nn} is independent of k. Such a network is called a neutral network [16]. If $k_{nn}(k)$ is an increasing function of k, then nodes with a high degree tend to connect to nodes with a high degree and the network is classified as assortative. If instead $k_{nn}(k)$ is a decreasing function of k, nodes with a high degree tend to connect to nodes with a low degree and the network is called disassortative [16].

However, the terms assortativity and disassortativity are also often used without taking degree correlations into account. Instead, one may look at a particular node property to classify networks as either assortative or disassortative [48]. In this way, an assortative network is a network where nodes tend to connect to other nodes with similar properties as themselves, whereas in a disassortative network nodes tend to connect to nodes with different properties [48].

2.1.5   Network distances

Several very useful measurements quantify the notion of length and distance in a network/graph. We start with the definition of a path: a path is a sequence of distinct nodes, such that adjacent nodes in the sequence are adjacent nodes in the network [26]. The length of a path is then defined as the number of edges in the path [26]. A geodesic path, often called the shortest path, between two nodes is the path between two nodes i and j with minimum length (note that the geodesic path is not necessarily unique) [16]. The length of the geodesic path between the two nodes i and j is called the geodesic distance or shortest distance d_{ij} [16]. Geodesic distance is often just called distance [26]. The notion of distance enables to define several network measurements such as average distance, diameter and radius.

Average distance

The average distance is the mean value of the geodesic distance d_{ij} [16]

$$l = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \ . \tag{2.11}$$

This definition, however, diverges if there are unconnected nodes in the network (unconnected nodes have, by definition, a distance equal to infinity) [16]. This problem could be avoided by only including connected pairs in the sum, but this introduces a distortion for networks with a lot of unconnected pairs of nodes. These networks will then have a small average distance, which is only expected for networks with lots of connections [16]. Instead, another definition of average distance is used: the global efficiency [16]

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \ . \tag{2.12}$$

This measurement quantifies the efficiency of the network in sending information between nodes [16]. This assumes that the efficiency of sending information between two nodes is reciprocal to their distance. The harmonic mean of the geodesic distances is now defined as the reciprocal of the global efficiency [16]

$$h = \frac{1}{E} \ . \tag{2.13}$$

Diameter and radius

The diameter $diam(G)$ of a network G is the maximum distance between any pair of nodes [26]. The eccentricity of a node is the maximum distance from that node to any other node. The radius $rad(G)$ is the minimum eccentricity among all nodes of G. Note that the diameter is the maximum eccentricity among all nodes in G [26].

2.1.6 Clustering coefficient and transitivity

The clustering coefficient is another important measure of network topology. It is a measure for the number of triangles in a network and determines the connectivity in the neighborhood of a node i: if a node i has a high clustering coefficient, its neighbors are likely to be directly connected to each other [37].

A triangle is defined as a loop of length three, this is a sequence of nodes x, y, z, x such that $\{x, y\}, \{y, z\}$ and $\{z, x\}$ are edges of the network. The clustering coefficient is thus a way to measure the degree to which nodes in a network tend to cluster [37]. The local clustering coefficient of a node i is defined as

$$C_i = \frac{2n_i}{k_i(k_i - 1)} \, , \tag{2.14}$$

where n_i is the number of edges that actually exist between the nodes in the neighborhood of i and $k_i(k_i - 1)/2$ is the maximum number of edges that could exist between them (note that this expression is only valid for an undirected graph, since for a directed graph $e_{ij} \neq e_{ji}$ and we have $k_i(k_i - 1)$ possible edges between the neighbors of node i). The average clustering coefficient is then given as the average of the local clustering coefficients

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \, , \tag{2.15}$$

where N denotes, once again, the number of nodes in the network. The clustering coefficient is, beside a measure for the connectivity in the network, also linked to the robustness, or resilience against random damage, of the network [28][33] [37].

A measure that is closely related to the clustering coefficient is the transitivity T, defined as (valid for undirected, unweighted networks) [16]

$$T = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triples of nodes in the network}} \, . \tag{2.16}$$

A connected triple is defined as a set of three nodes with at least two edges between them, so that each node can be reached from the other two (either directly or indirectly). The factor three arises from the fact that each triangle contributes to three different connected triples in the network: one centered at each node in the triangle [16].

Let us denote the number of triangles as N_Δ and the number of connected triples as N_3. These two numbers can be obtained from the adjacency matrix in the following way [16]

$$N_\Delta = \sum_{k>j>i} A_{ij} A_{ik} A_{jk} \, , \tag{2.17}$$

$$N_3 = \sum_{k>j>i} (A_{ij} A_{ik} + A_{ji} A_{jk} + A_{ki} A_{kj}) \, , \tag{2.18}$$

where A_{ij} are the elements of the adjacency matrix. It is possible to define the transitivity of one node as [16]

$$T_i = \frac{N_\Delta(i)}{N_3(i)} \ , \tag{2.19}$$

where $N_\Delta(i)$ represents the number of triangles that involve node i and $N_3(i)$ is the number of connected triples with i as the central node [16]

$$N_\Delta(i) = \sum_{k>j} A_{ij} A_{ik} A_{jk} \ , \tag{2.20}$$

$$N_3(i) = \sum_{k>j} A_{ij} A_{ik} \ . \tag{2.21}$$

It is not too hard to see that $N_\Delta(i)$ counts the number of edges between the neighbors of i and that $N_3(i)$ is equal to $k_i(k_i-1)/2$, where k_i is the degree of node i. It is then obvious that the local clustering coefficient C_i (Eq. (2.14)) of a node and the transitivity T_i (Eq. (2.19)) of a node define the same quantity [16]. In the case of the average clustering coefficient $\langle C \rangle$ and the global transitivity T this is, however, not the case. The difference between the two definitions is that the average of Equation (2.16) gives the same weight to each triangle in the network, whereas Equation (2.15) gives the same weight to each node. This may lead to slightly different values since nodes with a higher degree may possibly be involved in a higher number of triangles than nodes with a lower degree [16].

2.1.7   Network communities

Network communities are another feature that (complex) networks may posses. A community inside a network is loosely defined as a set of nodes that are more densely connected to nodes inside that set than to the other nodes in the network [46]. More strictly, a community is defined based on two hypotheses: the connectedness hypothesis and the density hypothesis [5]. In short, the connectedness hypothesis means that each member of a community should be reached through each other member of the same community [5]. The density hypothesis implies that nodes inside a community are more likely to be linked to other nodes inside that community than to nodes outside the community [5]. The density hypothesis narrows what could be considered a community, but it doesn't uniquely define it. Several community definitions are consistent with the density hypothesis. Let us consider three possible definitions: maximum cliques, strong communitoes and weak communities [5].

Maximum cliques

A clique is a complete subgraph, where a complete subgraph is defined as a set of nodes of the network where each node in the set is directly connected to all the others nodes in the same set [5]. A community based on this definition would then be the largest clique in the network. This definition of community might, however, be too restrictive and, beside that, large cliques don't appear very frequent in networks [5].

Strong communities

A strong community is defined such that each node inside the community has more edges to other nodes inside the same community than to nodes outside the community [5]. Let us denote a community as C and let i be a node inside the community C. If we define the internal degree $k_i^{\text{int}}(C)$ as the number of edges that connect node i with other nodes in C and the external degree $k_i^{\text{ext}}(C)$ as the number of edges that connect node i with nodes outside C, then we have the following condition for C to be a strong community [5]

$$k_i^{\text{int}}(C) > k_i^{\text{ext}}(C) , \qquad \forall i \in C . \tag{2.22}$$

Weak communities

A weak community is a community where the sum of the internal degree of all the nodes in the community exceeds the sum of the external degree of all the nodes in the community [5]. We thus have the following condition for a weak community C [5]

$$\sum_{i \in C} k_i^{\text{int}}(C) > \sum_{i \in C} k_i^{\text{ext}}(C) . \tag{2.23}$$

2.1.8 Modularity

The modularity Q is a measurement that represents the quality of a particular division of a network into different communities [16]. It was proposed after the fundamental problem arised for real networks concerning how to best divide the network into its constituent communities [16]. Generally, no a priori information is available about the number of existing communities in real networks [16].

The modularity for a network with N nodes and L edges and partitioned into n_c communities is calculated as follows. Each community has N_c nodes and L_c edges ($c = 1, ..., n_c$) [5]. The modularity of a community Q_c is then calculated by measuring the difference between the network real wiring diagram (which is given by the adjacency matrix A_{ij}) and the expected number of edges between i and j if the network is randomly wired (where i and j are two nodes that lie in the same community C_c) [5]. This last quantity is given by p_{ij}, which is obtained by randomizing the original network while keeping the expected degree of each node unchanged [5]. Hence, the modularity of a community Q_c becomes [5]

$$Q_c = \frac{1}{2L} \sum_{(i,j) \in C_c} (A_{ij} - p_{ij}) , \tag{2.24}$$

where the factor $\frac{1}{2L}$ is just a normalization factor. Using [5]

$$p_{ij} = \frac{k_i k_j}{2L} , \tag{2.25}$$

we can derive a simpler form for Q_c [5]

$$Q_c = \frac{L_c}{L} - \left(\frac{k_c}{2L}\right)^2 , \tag{2.26}$$

where L_c is the total number of edges in the community C_c and k_c is the total degree of the nodes in this community [5].

To generalize this idea to the whole network, we sum Eq. (2.26) over all n_c communities [5]

$$Q = \sum_{c=1}^{n_c} \left[\frac{L_c}{L} - \left(\frac{k_c}{2L}\right)^2\right] . \tag{2.27}$$

The derivation of this formula and thus of Eq. (2.26) can be found in Appendix A.1, Section A.1.1. The modularity Q can have values in the interval $[-1, 1]$ and has the following properties [5]

- The higher the value of Q is for a certain partition, the better is the corresponding community structure. This idea is used in community detection algorithms, where modularity is maximized in order to find the optimal community structure.

- A modularity Q equal to zero corresponds to the situation where all nodes are put in the same community, this is if the whole network is put into a single community. Then the two terms in Eq. (2.27) become equal and Q vanishes.

- If every node is put into a separate community, the first term in Eq. (2.27) vanishes and we are left with n_c negative terms. Hence, this situation corresponds to a negative modularity.

As said before, modularity is often used in community detection algorithms. It is thus important to be aware of the limitations of modularity. One of the biggest shortcomings of modularity maximization as community detection is its resolution limit. Modularity maximization forces small communities into larger ones [5]. If two communities A and B are merged, the modularity changes by [5]

$$\Delta Q_{AB} = \frac{l_{AB}}{L} - \frac{k_A k_B}{2L^2} , \tag{2.28}$$

where l_{AB} is the number of direct edges that connect nodes in community A with nodes in community B; k_A and k_B are the total degree of nodes in communities A and B respectively. The derivation of this formula is found in Appendix A.1, Section A.1.2. If A and B are distinct communities, they should remain distinct when Q is maximized [5]. This is however not always the case. Consider the case where $\frac{k_A k_B}{2L} < 1$ and there is at least one edge between A and B ($l_{AB} \geqslant 1$), then $\Delta Q_{AB} \geqslant 0$ and the modularity increases by merging the two communities [5]. Assume for simplicity that $k_A = k_B = k$, then modularity is increased by merging A and B if

$$k \leqslant \sqrt{2L} , \tag{2.29}$$

and if there is at least one edge between the two communities; even if A and B are distinct communities [5]! Equation (2.29) is called the resolution limit: communities that are smaller than this limit will not be detected by modularity maximization [5].

2.2 Network models

In order to study the topological structures of real-world networks, several theoretical network models have been developed. These theoretical models usually have a simpler representation and well known properties that can be derived analytically. They are widely studied and used extensively. The study of complex networks relies heavily on the knowledge and understanding of these models. In this section we will review some of the most important ones.

2.2.1 The Erdős-Rényi model

In 1959 the two mathematicians Paul Erdős and Alfréd Rényi proposed a model to generate simple random networks [16]. Independently of them, Solomonoff and Rapoport already proposed the model in 1951. The model constructs random networks with N nodes in the following way [3]

1. Start with N disconnected nodes,

2. For each pair of nodes, connect them with a predefined probability p.

If $p = 1$, we obtain a network with the maximum number of edges $N(N-1)/2$. This is called the fully connected network.

Degree distribution and average degree

The degree distribution of the random network with N nodes is given by

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} . \tag{2.30}$$

The derivation can be found in Appendix B.1, Section B.1.1. From this we can calculate the average degree of the network. The expected mean degree is given by [52]

$$\begin{aligned}
\langle k \rangle &= \sum_{k=0}^{N-1} k\, p_k \\
&= \sum_{k=0}^{N-1} k \binom{N-1}{k} p^k (1-p)^{N-1-k} .
\end{aligned} \tag{2.31}$$

This equation can be simplified to $\langle k \rangle = (N-1)p$, and for $N \gg 1$ this approximately becomes $\langle k \rangle = Np$. The derivation of this result can be found in Appendix B.1, Section B.1.2.

Most real networks are sparse, which means that $\langle k \rangle \ll N$. In this case the degree distribution of the random network is well approximated by a Poisson distribution [4]

$$p_k = \mathrm{e}^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} . \tag{2.32}$$

The derivation can be found in Appendix B.1, Section B.1.3.

Average clustering coefficient

The average clustering coefficient (Eq. (2.16)) of the Erdős-Rényi network is given by [11]

$$\langle C \rangle = \frac{\binom{N}{3} p^3}{\binom{N}{3} p^2} = p = \frac{\langle k \rangle}{N - 1} \; . \tag{2.33}$$

This can be easily understood since choosing three nodes out of N is given by $\binom{N}{3}$. The chance that they are connected by three edges is given by p^3, which gives the number of triangles; the chance that the are connected by two edges is p^2, which determines the number of connected triples (recall that this model creates random networks with no self-loops and no edge duplicates). If the average degree $\langle k \rangle$ is constant, the average clustering constant goes as $\langle C \rangle \sim \mathcal{O}(1/N)$, which becomes very small for large networks [11].

Diameter and average distance

A heuristic derivation of the diameter of a random network goes as follows. Consider a random network with average degree $\langle k \rangle$, then a node in this network has on average [4]

- $\langle k \rangle$ nodes at distance one ($d = 1$),

- $\langle k \rangle^2$ nodes at distance two ($d = 2$),

- $\langle k \rangle^3$ nodes at distance three ($d = 3$),

- ...,

- $\langle k \rangle^d$ nodes at distance d.

The number of nodes at a distance d cannot exceed the number of nodes N in the network. We thus obtain the following equation from which the maximum distance or diameter $d_{\max}$ can be obtained

$$N \approx \langle k \rangle^{d_{\max}} \; , \tag{2.34}$$

or

$$d_{\max} \approx \frac{\ln N}{\ln \langle k \rangle} \; . \tag{2.35}$$

A more precise calculation can be found in Appendix B.1, Section B.1.4. Equation (2.35) might, however, be a better approximation to the average distance $\langle d \rangle$ between two randomly chosen nodes [4]. This is because $d_{\max}$ is often dominated by a few extreme paths, where $\langle d \rangle$ instead is averaged over all node pairs and thus suppresses these fluctuations [4]. Hence, the average distance of a random network is defined by [4]

$$\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle} \; . \tag{2.36}$$

Since for large N, $\ln N \ll N$, we find that the average distance is a lot smaller than the size of the network. Generally, the logarithmic dependence of the average distance $\langle d \rangle$ on the network size N is one of the two defining properties of what is referred to as the "small world phenomenon"; the other property is a high clustering coefficient [20].

To summarize: the random Erdős-Rényi network has a bell shaped degree distribution which approximates a Poisson distribution for $N \gg 1$; the average clustering coefficient becomes small for large networks and the average distance depends logarithmic on the system sizes. The Erdős-Rényi network thus possesses one of the two properties of the "small world phenomenon" (namely short average distance).

Real social networks, however, usually display the "small world phenomenon" and thus have both a short average distance and a high clustering coefficient. The question then rises: can we construct a model that has both of these properties? The answer is yes and this is exactly what Duncan Watts and Steve Strogatz did.

2.2.2 The Watts-Strogatz model

In 1998, Duncan Watts and Steve Strogatz constructed a network model that generates random networks with small world properties, such as high clustering and short average distances. Their model achieves this by interpolating between a regular ring lattice and a random network [50]. The construction goes as follows [50]

1. Start with a regular ring lattice with N nodes, each connected to $K = 2m$ neighbors (m on each side).

2. For each node i, rewire each edge that connects i with its $K/2$ rightmost neighbors with a probability β. Rewiring is done in such a way that self-loops and edge duplication are avoided.

The probability β allows us to tune the graph between a regular lattice ($\beta = 0$) and a random network ($\beta = 1$) [50]. Usually, the following condition is required for the number of neighbors K: $N \gg K \gg \ln N \gg 1$. The first inequality assures sparse networks, whereas the second inequality guarantees that a random graph will be connected (thus the network is not so sparse that it becomes disconnected) [50].

The regular ring lattice ($\beta = 0$) has an average clustering coefficient $\langle C(0) \rangle$ and an average path length $\langle d(0) \rangle$ equal to [50]

$$\langle C(0) \rangle = \frac{3(K-2)}{4(K-1)} \approx \frac{3}{4} \text{ for } K \gg 1 , \qquad\qquad \langle d(0) \rangle \approx \frac{N}{2K} . \qquad (2.37)$$

The derivations of these formulas can be found in Appendix B.2, Section B.2.1. The regular ring lattice thus has a high clustering coefficient and an average path length that grows linearly with the system size N.

The random network ($\beta = 1$) on the other hand has a low clustering coefficient $\langle C(1) \rangle$ and a small average path length $\langle d(1) \rangle$ (see Eq. (2.33) and Eq. (2.36)) [50]

$$\langle C(1) \rangle \approx \frac{K}{N} \ll 1 , \qquad\qquad \langle d(1) \rangle \approx \frac{\ln N}{\ln K} . \qquad (2.38)$$

Watts and Strogatz then showed by numerical simulation that the average distance for the networks with $0 < \beta < 1$ is comparable with the average distance of a random graph (which is small), even for small β [40]. The clustering coefficient is approximately given by [6]

$$\langle C \rangle \approx \langle C(0) \rangle \left(1 - \beta\right)^3 , \tag{2.39}$$

where $\langle C(0) \rangle$ is given by Equation (2.37). This can be seen in the following way

- The probability that none of the three edges of the original triangles in the regular lattice are rewired is $(1 - \beta)^3$.

- The probability that edges are rewired back to each other is negligible.

For rather small values of β, the clustering coefficient remains high. The model of Watts and Strogatz is thus able to generate random networks with a high clustering coefficient and short average distances.

2.2.3 The stochastic block model

The stochastic block model (SBM) is another random graph model that is able to produce networks containing a community structure. The stochastic block model originates in social sciences and was developed to describe group structures in friendship networks [25]. The idea behind the development of this model is that often in real networks, there are groups of nodes that are more densely inter-connected with each other than with the rest of the network [16]. Hence, these networks posses a modular/community structure and this is exactly the structure that is easily modeled by the stochastic block model. The stochastic block model is also often used as a benchmark for community detection algorithms [2].

The stochastic block model in its most simple form is defined by the following parameters [12]

- The number of nodes N,

- A partition of the node set $\{1, 2, ..., N\}$ into k disjoint subsets $C_1, ..., C_k$; these groups are the communities,

- A symmetric $k \times k$ matrix M; the diagonal elements M_{ii}, $i = 1, ..., k$, represent the edge prob-abilities inside the community C_i and the off-diagonal elements M_{ij}, $i, j = 1, ..., k$ with $i \neq j$, give the edge probabilities between nodes in community C_i and nodes in community C_j.

There are no real restrictions on the edge probabilities, which makes the stochastic block model a very general model that is capable of reproducing many different network structures [34]. However, the edge probabilities inside a community are often taken to be bigger than the probabilities between com-munities ($M_{ii} > M_{ij}$, $i \neq j$). In this way, the stochastic block model generates a network with a clear community structure. A stochastic block model that fulfills this condition is called an assortative SBM [36]. A disassortative stochastic block model refers to the case where the edge probabilities between the communities are higher than those inside the communities ($M_{ii} < M_{ij}$, $i \neq j$) [27].

The stochastic block model includes the Erdős-Rényi model [12]. This can be seen in the two different

ways. First, consider the case where $k = 1$, thus we only have one community and all nodes belong to this community. The matrix M then reduces to one number p and we are left with the parameters N and p of the Erdős-Rényi model [12]. Second, assume we have more than one group or community, $k > 1$, and assume that every element of the matrix M is the same, $M_{ij} = p, \forall\, i, j$ [12]. Again, we are left with the parameters N and p of the Erdős-Rényi model. In both cases, the mathematical properties of the Erdős-Rényi model are obtained: a Poisson degree distribution, an average distance that is logarithmic with the system size N, vanishing clustering coefficient for large N, etc [12]. The structure within the groups or communities of the stochastic block model is that of a random graph, while the structure between the communities is that of a random bipartite graph [12].

In this thesis, the stochastic block model is generated in the following way. First, the parameters of the model are specified. Then, for each community C_i an Erdős-Rényi network is constructed, where the number of nodes equals the size of the community C_i and the edge probability p is given by M_{ii}. Finally, edges are added between communities C_i and C_j with probability $p = M_{ij}$.

The stochastic block model can be fitted to data [12]. Given a choice of the number of communities k and an observed network G, the stochastic block model can be used to infer the latent community structure and the edge probability matrix M [12].

One of the shortcomings of fitting the SBM to infer the community structure, is that real-world networks often have a heterogeneous or power-law degree distribution. Since the SBM is made of random networks inside each group and random bipartite networks between each pair of groups, the degree distribution of the full network is always a mixture of Poisson distributions [12]. The consequence of this is that if a network possesses a more skewed degree distribution, then fitting the SBM to the data will most likely give poor results [34]. The reason behind this is one of the main assumptions of the stochastic block model, namely the fact that the edge probabilities only depend on which group the node belongs to [12]. In other words, nodes that belong to the same group are stochastically equivalent, meaning that they have equivalent connectivity patterns to other nodes [12]. However, a network with a skewed degree distribution violates this assumption of edge independence [12]. To say it in the words of Karrer and Newman: "Just as the fitting of a straight line to intrinsically curved data is likely to miss important features of the data, so a fit of the simple stochastic block model to the structure of a complex network is likely to miss much [...]." (Karrer and Newman, p.1, reference [34]). Karrer and Newman proposed a variation of the SBM, the degree corrected SBM, that deals with this limitation. We will not discuss their variation, since that would lead us too far astray, but interested readers can find more detailed information in reference [34].

An extension that we are going to explore a bit deeper is the stochastic Watts-Strogatz block model (SBM-WS).

The stochastic Watts-Strogatz block model (SBM-WS)

As said before, the stochastic block model is a very flexible model and there are many different extensions. One extension that is used in this thesis is a stochastic block model that uses the Watts-Strogatz model instead of the Erdős-Rényi model to construct the constituent communities. In this way, we obtain a version of the stochastic block model that has both a community structure and a high clustering coefficient.

The construction of the model is very similar to that of the regular stochastic block model. For a SBM-WS with n_c communities of specified sizes, we first construct n_c Watts-Strogatz networks (for each constituent Watts-Strogatz network the rewire probability β and the mean degree K are specified), then edges are (with some predefined probability) randomly added between the different Watts-Strogatz communities. The result is a network that can be tuned to have both a community structure and a high clustering coefficient. This is different from the regular stochastic block model that can only account for a community structure and lacks a high clustering.

2.3   Real-world social networks

The models discussed in the sections above, describe real-world networks in an idealized, simplified way. They may posses some of the characteristics of real-world networks, but they certainly do not always describe real-world networks appropriately. The question then arises: what are the main properties/characteristics of real-world networks, and more specifically of real-world social networks? Another important question that can be asked is whether or not there is a significant difference between off-line social networks and on-line social networks?

As already briefly discussed in the introduction, it is found that real-world networks often posses three characteristic properties. These are the two properties of the "small world phenomenon", namely a short average distance and a high clustering coefficient, and a heterogeneous or power-law degree distribution [13]. These properties are valid for both off-line and on-line social networks [51]. Lets discuss them in a bit more detail.

Short average distance

In the late 1960's Stanley Milgram performed a, what is now famous, experiment. He asked randomly selected people in Boston and Omaha to forward a letter to a distant target person. The letter could, however, only be sent to acquaintances, thought to be closer to the target person [13]. The remarkable result of this experiment was that the average number of steps between the sender and the target person was only around six. This phenomenon is now often referred to as "six degrees of separation" [13]. This phenomenon is also observed in many other real-world networks: in most real-world networks it is possible to go from one node to another node through a number of edges that is small compared to the system size [13]. Such networks usually have an average distance that depends logarithmically on the system size instead of linearly. A short average distance facilitates a fast transmission of information [51].

High clustering coefficient

In real-world social networks, one often observes the pattern that two acquaintances/friends of an individual are also acquaintances/friends of each other [13]. In graph terminology, this is translated into a high clustering coefficient (which quantifies a large number of triangles in the network) [13]. This property is not only found in social networks, but also in a wide variety of different real-world networks [13].

Heterogeneous or power-law degree distribution

Many real-world networks are found to have a power-law degree distribution [13]

$$P(k) \sim k^{-\gamma} \,, \tag{2.40}$$

where the exponent γ often lies between 2 and 3. Networks with a power-law degree distribution are also called heterogeneous or scale-free networks. The term scale-free refers to the fact that there is no characteristic scale in the network. Compared to the homogeneous network, where each node has a degree close to the average degree $\langle k \rangle$, the heterogeneous network has many nodes with a low degree and a few nodes with a very high degree (these nodes are referred to as hubs). If the exponent γ is smaller or equal to two ($\gamma \leqslant 2$), the average degree diverges; if $2 < \gamma < 3$, there is a finite average degree, but a diverging variance [41].

As said earlier, these three main characteristics are valid for on- and off-line social networks. Another property that is found in most on- and off-line networks is the appearance of a giant component, this is, most nodes belong to the same connected component [35]. They also often posses a modular/-community structure [22][38].
There are, however, some properties that are different between on-line and off-line social networks as well. One of them might be the degree correlations/mixing pattern. It is commonly assumed and accepted that most real-life social networks have an assortative mixing pattern (this distinguishes them from other real-world networks, eg. biological networks [16], which often have a disassortative pattern) [31][51]. For a long time, it was thought that this is also valid for on-line social networks. However, recent research has shown that a lot of on-line social networks display disassortative mixture [31][51]. This discrepancy between real-life social networks and on-line social networks may be explained as follows. In real life, 'ordinary people' often want to be friends with celebrities, however, the celebrities usually prefer to stay within their own circles [31][51]. The 'ordinary people' don't have access to these circles and we are left with the situation where 'ordinary people' interact with 'ordinary people' and celebrities hang out among other celebrities; hence, an assortative structure. In on-line social networks, 'ordinary people' can easily connect with the celebrity and vice-versa, the celebrity wants to show his/her importance/influence by the number of fans [31][51]. This leads to a disassortative mixing pattern.

Chapter 3

Opinion dynamics

The study of opinion dynamics in a group of people, as already mentioned in the introduction, has two major layers: a network layer and an opinion dynamics layer. The network layer, which describes the underlying structure and correlations between the agents/people, is intensively described in the previous chapter, Chapter 2. The current chapter focuses on the second important layer, namely that of the opinion dynamics models. These models need to describe, in a simplified way, how opinions are formed and how they evolve in a group of people that interact in a certain way. Hence, these opinion dynamics are the dynamics that run on the underlying network structure.

Several models have been proposed; here we will review some of the most important ones, with the main focus on the models used in this thesis.

Some major attention will also be given to the content curation/filtering algorithms used by social media companies (see Section 3.3). This is needed because this thesis deals with opinion dynamics on on-line social media, so the aspect of filtering cannot be ignored in the formation and evolution of opinions.

There will also be some attention on the activation mechanism (see Section 3.4). This is the mechanism that describes the timings in which content/information is exchanged between users [45]. Such a mechanism may change the network from a static one to a more temporal one, where nodes become active and inactive over time.

There are two major groups of opinion dynamics models, namely binary/discrete models and continuous approaches. These will be discussed in Section 3.1 and Section 3.2, respectively. This thesis only deals with binary opinion dynamics models, so these will be thoroughly described; the continuous models will be quickly reviewed in order to give a more complete overview.

Finally, in Section 3.5, the introduction of stubborn people is discussed.

3.1 Binary models

Binary models are models where each agent can have one of the two opinions A or B (or, equivalently, 0 or 1). A binary model could reflect real-life cases where there exist two competing opinions, such as a vote in an election system with two competing candidates or choosing between Windows and Linux [42].

In this section, some binary models will be discussed; note, however, that the models described here are far from all the models that currently exist.

3.1.1 The majority model

One of the simplest opinion dynamics models is the binary majority model. In this model each agent has one of the two opinions A or B (or 0 or 1). Initially, these opinions can be equally distributed or not. If the distribution is not equal, but instead eg. 80/20, there is a predominant opinion and a minority opinion.

The following updating rule is then imposed on the network: each agent/individual adopts the majority opinion of his/her nearest neighbors [42]. Thus, initially each agent is assigned an opinion; then the opinions are evolved in time: each individual iteratively updates his/her opinion according to the majority opinion of his/her direct neighbors [42].

Several studies have been performed on this kind of model, searching for answers on several questions such as: what are the conditions that drive the system to converge to unanimity; and, does an opinion, that is initially a majority opinion, remain a majority opinion in the final state [42]?

Most of these studies conclude that, at the end of the opinion evolution, a complete consensus state is obtained where all the agents adopt the same opinion [42]. This is even the case when the opinions are equally and randomly distributed in an Erdős-Rényi network (with the side note that the degree needs to be high enough), as was shown by Benjamini, Chan, O'Donnell, Tamuz and Tan [7] (also note that even when the opinions are equally distributed, there may still be a slight imbalance due to the random assignement of an opinion to each node; instead of a perfect 50/50 distribution, one may end up with a eg. 49.8/50.2 distribution; in this sense there is still some kind of notion of 'minority' and 'majority' opinion). It is also found that the rate of convergence increases with increasing edge density and that if the network is sparse enough, consensus will not be reached and the two opinions coexist [42]. This can be explained by the fact that if the network is sparse enough, we may have disconnected parts; each of these disconnected groups may then evolve to a different consensus opinion, so that over the entire network different opinions coexist [42].

3.1.2 The probabilistic majority model

The probabilistic majority model is similar to the majority model, but has, as the name suggests, a more probabilistic/random nature. In this model the fraction of neighbors of user i with opinion A and the fraction of neighbors with opinion B are determined. Then, the opinion of user i is updated: he/she takes on opinion A (or B) with a probability equal to the fraction of his/her neighbors with opinion A (or B) [45].

For example, if 70% of the neighbors of user i have an opinion A (and thus 30% of the neighbors have opinion B), user i will update his/her opinion to opinion A with a probability 0.7 and to opinion B with a probability 0.3 [45]. Thus, user i no longer simply updates his/her opinion according to the majority of opinions carried by his/her neighbors, but the majority opinion still has a higher probability to be taken on by this user.

This model is no longer deterministic, but has a level of randomness, which is an important component of opinion dynamics [45].

3.1.3   The voter model

Another extremely simple binary model, maybe even simpler than the majority model, is the voter model. The model works as follows: each agent in the network has a binary opinion (A or B); at each time step, an agent is randomly selected along with one of its neighbors and, subsequently, the agent adopts the opinion of this neighbor [10]. The agent thus mimics its neighbor [10].

The pressure of the majority in only felt in an averaged way, in the sense that a neighbor that carries the majority opinion has more chance to be randomly selected [10].

Lots of studies have been performed on this model. In this thesis, however, we will not make use of the voter model. Hence, we will not go into further detail. Interested readers are referred to references [10], [21] and related references therein.

3.2   Continuous models

In continuous opinion dynamics models, the opinion variable of the agents in the network is no longer discrete, in contrast to binary or discrete models, but takes real values (often in the interval $[0, 1]$ or $[-1, 1]$).

In continuous opinion dynamics models, some concepts that are valid for the discrete models no longer apply (eg. the concept of a majority/minority opinion or the concept of equality of opinions); hence, they require a bit of a different framework [10].

The basis of most continuous models is a set of N nodes or agents, each having an opinion x_i represented by a real number that takes values in some interval [10]. In contrast to discrete/binary models, all agents usually start with different opinions [10]. The possible outcomes are more complex than in the discrete models and often has opinion clusters emerging in the final states (where opinion clusters refer to clusters of nodes with the same opinion) [10]. The number of opinion clusters can be one (consensus), two (polarization) or more (fragmentation) [10].

Several models have been proposed, we will only discuss a fraction of them.

3.2.1   The classical models

In the classical model, there are N nodes/agents, each with an opinion x_i and each node i gives a weight w_{ij} to all its neighbors (with $w_{ij} \geqslant 0$; this weight may for example represent the importance/influence of the neighbors) [29]. The most simple model has fixed weights w and the updating rule is [29]

$$x(t+1) = Wx(t)\,,\tag{3.1}$$

where W is the fixed stochastic matrix which contains the weights w_{ij}, $x(t)$ is the opinion profile at time t (this is a N-dimensional vector containing the opinions x_i of each node at time t) and $x(t+1)$ is the opinion profile at time $t+1$.

Friedkin and Johnsen proposed a variation of this model [23][24]. The model they proposed assumes that an agent i adheres to its initial opinion to a certain degree g_i and is influenced by others by a susceptibility of $1 - g_i$ [29]. This variation thus has the following updating rule (in matrix notation) [29]

$$x(t+1) = Gx(0) + (\mathbb{1} - G)Wx(t) \, , \tag{3.2}$$

where G is the diagonal matrix which contains the g_i and $\mathbb{1}$ is the identity matrix. If all the g_i are equal to zero, we obtain the model of Eq. (3.1) [29].

A possible time variant extension is [29]

$$x(t+1) = W(t)x(t) \, , \tag{3.3}$$

where now the weights are time dependent. This model can be used for situations in which so called "hardening of positions" occur [29]. This means that, over time, the agents put more and more weight on their own opinion and less weight on the opinion of others [29].

3.2.2   Bounded confidence models

In bounded confidence models, each node has, beside an opinion x_i, a threshold ϵ_i, also called an uncertainty or tolerance (note that ϵ can be the same for all nodes or can be different for each node) [10][18]. This threshold determines which opinions of neighbors are close enough to the agent's opinion in order to influence that agent; opinions that are beyond this threshold are ignored [18]. In other words, an agent with opinion x_i is only affected/influenced by those neighbors which have an opinion in the interval $[x_i - \epsilon, x_i + \epsilon]$ [10].

The two most popular bounded confidence models are the Deffuant model and the Hegselmann-Krause model [10].

The Deffuant model

Assume we have a population of N agents, each with an opinion x_i that takes real values in the interval $[0, 1]$ [10]. Furthermore, the threshold ϵ is the same for all the agents. At each time step, an agent i is randomly selected together with one of its neighbors j (which is also randomly selected) [10]. If the opinions of the two nodes i and j are too far apart nothing happens; this is the case if $|x_i(t) - x_j(t)| \geqslant \epsilon$. If instead $|x_i(t) - x_j(t)| < \epsilon$, then the opinions are updated to [10]

$$\begin{aligned} x_i(t+1) &= x_i(t) + \mu[x_j(t) - x_i(t)] \, , \\ x_j(t+1) &= x_j(t) + \mu[x_i(t) - x_j(t)] \, . \end{aligned} \tag{3.4}$$

The parameter μ is called the convergence parameter; its value lies in the interval $[0, 1/2]$ [10]. The model is based on a compromise strategy: after some discussion and a constructive debate, the

opinions of the two agents move closer to each other [10].

The final state is a situation where several opinion clusters coexist. This is because the opinions move closer and closer to each other in several clusters; once each cluster is sufficiently far apart from the others (so the difference in opinions between all agents in the two clusters is larger than ϵ), they no longer interact with each other [10]. Only agents inside the same cluster keep on interacting with each other, which makes the opinions inside a cluster converge to a single value [10]. Usually, the number and sizes of the clusters depends on the threshold ϵ, whereas μ determines the convergence rate [10].

The Hegselmann-Krause model

The Hegselmann-Krause model is very similar to the Deffuant model. Again, there are N agents, each with an opinion x_i and a constant threshold ϵ [10]. Each agent interacts with those neighbors with an opinion in the interval $[x_i - \epsilon, x_i + \epsilon]$ [10]. The difference with the Deffuant model lies in the updating rule [10]. In the Deffuant model, only the opinion of one randomly selected neighbor is taken into account, whereas the Hegselmann-Krause model considers all compatible neighbors [10]. The updating rule becomes [10]

$$x_i(t+1) = \frac{\sum_{j:|x_i(t)-x_j(t)|<\epsilon} A_{ij} x_j(t)}{\sum_{j:|x_i(t)-x_j(t)|<\epsilon} A_{ij}} \ , \tag{3.5}$$

where A_{ij} are the elements of the adjacency matrix. This model thus takes the average opinion of the compatible neighbors into account [10].

The dynamics of this model is similar to the Deffuant model and the final state is also characterized by a number of stationary opinion clusters [10]. The number of clusters decreases if ϵ increased and above a critical value ϵ_c of the threshold, there can only be one cluster [10].

There are, of course, many other variations of the bounded confidence model. As mentioned earlier, the threshold ϵ can be different for each node, so that each node has an opinion x_i and a threshold ϵ_i or the confidence interval can be made asymmetric $[x_i - \epsilon_l, x_i + \epsilon_r]$ with $\epsilon_l \neq \epsilon_r$ [29].

3.3 Algorithmic personalization

On-line social platforms use algorithms to order and filter the posts that appear on the time-line of an individual [45]. They do this in order to make the experience on their platforms more pleasant, interesting and convenient [45]. If social media companies wouldn't use any kind of filtering, there would be way to many content to handle and process for the social media user [8]. This in turn would make the on-line social experience a lot more unpleasant and inconvenient; the user might feel overwhelmed by the number of information he/she is exposed to [8]. The reason behind this is that people are cognitive and time constrained [45]; hence, there is only a limited amount of information that humans can process [8].

The backside of such algorithmic filtering is that it fine-tunes the information it shows based on the user. Even two users with the same friends may then be exposed to different content, based on their

previous activities and history [8]. This may lead to the formation of echo chambers in which users get 'trapped' in content that confirms their own mindset [8].

It is thus important to know and be aware of the possible impacts of algorithmic personalization on the formation and evolution of opinions. Prof. Nicola Perra and prof. Luis E.C. Rocha already partly investigated the role of such algorithms on the dynamics of opinions in several networks in one of their previous papers, *"Modelling opinion dynamics in the age of algorithmic personalization"* (see reference [45]). Since filtering algorithms are often not transparently shared by social media companies, they proposed several personalization algorithms based on three simple ways of filtering [45]

- Popularity filtering: promoting content that is popular across the platform,

- Semantic filtering: promoting posts that are similar to previous consumed posts,

- Collaborative filtering: promoting content that people similar or close to us consume.

The following paragraph briefly reviews the filtering algorithms that prof. Nicola Perra and prof. Luis E.C. Rocha proposed.

At every time t each platform user has a hidden list $L_i(t)$ that contains all the opinions/posts that he/she received from his/her friends [45]. The filtering algorithm then selects S items from this list and shows them on the user's time-line (this is the fundamental idea behind algorithmic personalization: only showing a fraction of the possible posts [45]); the following selection methods are used [45]

- The reference method (REF): this method randomly selects S items from the list $L_i(t)$.

- The old method (OLD): the S oldest post in $L_i(t)$ are displayed on the time-line of the user.

- The recent method (REC): shows the S most recent items from $L_i(t)$.

- The preference method (PR): sorts the hidden list $L_i(t)$ according to the current opinion of the user and then selects the first S elements. This filtering method is based on the semantic filtering explained above.

Prof. Nicola Perra and prof. Luis E.C. Rocha showed that the REF, OLD and REC methods behave similar in many cases and that the PR method often displays different behavior (such as a larger formation of echo chambers compared to the other three methods) [45]. In this thesis the REC and REF methods are mainly used as a reference/benchmark method and the PR method is used as the main filtering algorithm.

3.4 The activation mechanism: Static to temporal

In the most simple opinion dynamics models, which make use of a static network, users are always active and thus always exchanging opinions/posts with each other. In this kind of model the opinion dynamics might go as follows: each time step a random user is selected after which his/her opinion is updated according to the opinion dynamics model that is implemented. Another possibility is updating all users at the same time during each time step.

It might, however, be a more realistic situation if nodes/users become active/inactive over time. This

would be a better representation of real social media platforms, where users are active for some periods of time and then become inactive again. This kind of activation mechanism of the users gives the network a more temporal character. The underlying network structure doesn't change (in the sense that each user/node has a fixed number of neighbors that doesn't change over time, so no new friends are made or, conversely, no old friends are being defriended), but by making users active and inactive the exchange of information between users becomes time dependent and depends on which users are active and which aren't. Hence, in this sense, this does affect the network structure in some way, because if a user, for example, doesn't become active for a long period of time, he/she doesn't participate in the exchange of information/opinions and the network structure behaves like if this user/node is absent. This may affect the dynamics, since it may seem as if the network has a different average degree and some properties like connectivity, clustering coefficient, etc may behave differently.

In this thesis we follow the same activation mechanism as was proposed in the paper of prof. Nicola Perra and prof. Luis E.C. Rocha (see reference [45]). Every time step users are made active with some probability p_i (thus, every time step, each node becomes active with a probability p_i). In this thesis the activation probability is set equal to $p = 0.1$, which is constant for all the nodes. Prof. Nicola Perra and prof. Luis E.C. Rocha chose, in their paper, activation probabilities that are extracted from a power-law ($F(p) \sim p^{-1.5}$, with $p \in [0.01, 1]$) [45]. They did this based on the analyses of real social networks that shows that the propensity of users to start an on-line social interaction follows a heterogeneous distribution [44][45]. However, they also showed that using a constant activation probability p over the entire population yields results that are qualitatively similar (at least in their case) [45].
Hence, for simplicity, this thesis implements the constant activation probability variant. It might however be an interesting extension to switch to the heterogeneous activation probabilities and investigate the possible impacts on the dynamics.

If a user is active, he/she first spreads his current opinion to all his/her neighbors (for example by posting something). Then, the active user updates his own opinion according to the opinions on his/her time-line [45]. This updating goes according to the implemented opinion dynamics model (eg. majority model, probabilistic majority model,...). In every time step the updating of all active nodes is done simultaneously [45].
It is important to note that what somebody sees on his/her time-line is the consequence of the interplay between the filtering algorithm and the posted opinions of his/her neighbors. Every user has a "hidden time-line $L_i(t)$" which contains the actual time-line of the user at the moment when he/she was active for the last time (t_{last}) plus all the posts of the neighbors from that last time he/she was active till the current time he/she becomes active again (t) [45]. At the time t that the user becomes active again, the filtering algorithm selects Q post from this "hidden time-line $L_i(t)$" and shows them on the actual time-line of the user; these Q items are thus the opinions that the user is exposed to [45] (see also Section 3.3 for more detailed information on the filtering algorithms).

3.5   The presence of stubborn people

It is well known that in real life, you can find people that are really persistent towards their own opinion and that are not open to opinions of others. They are convinced that they carry the truth with them and that other people, with different world-views, are not capable of giving them any new and useful insights. They are close-minded in contrast to people with an open-mind that are willing to learn from others and to reconsider their beliefs. Of course this is put a bit simplistic and there are many shades of gray between this black and white statement. However, there are groups in society, often with more extreme opinions, that are know to be really inflexible towards changing that (extreme) opinion (note here that being stubborn and being extreme doesn't need to go hand in hand, you can obviously be stubborn without being extreme and vice versa; however, it is often widely accepted that there is some kind of correlation/link between people with extreme opinions and being more persistent towards these opinions compared to people with more moderate beliefs; this idea of correlating extreme people with a more stubborn character is also often used in models, see for example [18][49]) (also note that having some resistance to change your opinion to any other opinion doesn't necessarily need to be a bad thing, it can be seen as a form of critical thinking and as a form of taking things with a grain of salt).

Knowing the impact of such stubborn groups on the opinions of others is of great importance, especially with the rise of social media which makes the reach of such extreme and persistent groups a lot bigger. Before the era of social media, one could maybe argue that these extreme and stubborn groups are often isolated and therefore wouldn't reach too many people. So their impact could only be of a minor nature (note that this is only a statement that could have been given, there are of course plenty of examples throughout human history that contradict this argument; to only mention a few: Nazism in the 1930s and 1940s, the Ku Klux Klan in the late 19th and begin 20th century, the rise of communism in the 20th century, etc, all examples of groups that can be considered extreme and that have had a wide reach and a significant impact on society). Today, this possible argument is no longer valid; social media enables the contact with all sorts of beliefs and all sorts of people with almost no (time and spatial) constraints. Extreme and stubborn groups that may have been isolated before no longer need to be so and this may have huge, and largely unknown, impacts.

What is the impact of having stubborn/extreme people in your society? Does it matter if the stubborn people are grouped together or if they are distributed evenly throughout your system? Are there ways to reduce the possible impacts of stubbornness and/or extremism? Can these stubborn groups lead to polarization or even overtake the opinions in the system? Or is introducing some kind of stubbornness, on the other hand, a protective factor towards extremism? To answer these questions, several models have been proposed to include some kind of stubbornness/resistance in the system under study.

3.5.1   Stubbornness in bounded confidence models

The bounded confidence models, discussed in Section 3.2.2, can be easily extended to included some notion of stubbornness. Guillaume Deffuant proposed such an extension where stubbornness is correlated to extremism [18]. The model he proposed goes as follows. Remember that in the bounded confidence models, the agents have an opinion x_i (which is a real number, often between [0, 1] or between [−1, 1]) and a threshold ϵ_i, which determines the range of opinions that can influence agent i (agents with opin-

ions between $[x_i - \epsilon_i, x_i + \epsilon_i]$ are capable of influencing agent i). Guillaume Deffuant then introduced a fraction of extremists, f_e, which can have either the extreme opinion -1 or the extreme opinion 1 [18]. These extremists are given a threshold equal to ϵ_e [18]. The other agents are considered to be moderate and can have opinions between -1 and 1. The moderates are given a threshold ϵ, with $\epsilon > \epsilon_e$ [18]. This last condition implicitly denotes the resistance of the extremes towards other opinions in contrast to the more flexible nature of the moderates, which is again based on the idea that people with extreme opinions will be more convinced of that opinion and will be less likely to change it.

Since this thesis is not concerned with bounded confidence models, we will not discuss the results of this model here. Interested readers can find them in reference [18].

Another example of introducing some kind of persistence towards the own opinion of the nodes is the "hardening of positions" that was briefly discussed at the end of Section 3.2.1.

3.5.2 Stubbornness in binary models

Playing around with the threshold ϵ_i of the agents to include a notion of stubbornness, such as one can do for the bounded confidence models, is obviously not possible for binary/discrete models. There are, however, many other ways to include some kind of stubborn people in these models. In the following, two different possibilities will be discussed.

One small remark must be made here: previously, the link between stubbornness and extremism was often made and it was often assumed that the more extreme somebodies opinions are, the more persistent that somebody is towards those opinions. In binary models, however, there are only two possible opinions that can be taken on and in that sense it is less useful to talk about extreme and moderate opinions. However, we can still have people that rather keep their own opinion than be convinced by friends to switch to the other opinion. Whether this resistance is a sign of critical thinking or ignorance/being convinced that you know it best, is up to the reader to decide. The models simply introduce some type of resistance and make no distinction between possible underlying reasons for this resistance.

Including a stubbornness parameter

One way of creating stubborn people in the binary model is by giving each node i a stubbornness/resistance parameter r_i. This resistance parameter can have values between 0 and 1, $0 \leqslant r_i \leqslant 1$.

- If $r_i = 0$, the node is not stubborn and can be easily convinced of the value of the other opinion (according to which opinion dynamics model that is implemented: majority, probabilistic majority,...).

- If $r_i = 1$, the node is completely stubborn and there is no chance of convincing him/here of the other idea.

- If $0 < r_i < 1$, the node has some resistance but can still be convinced to switch to the other opinion, even though he/she will be less likely to do so. A uniform random number f between 0 and 1 is drawn. If $f > r_i$, the node will update his/her opinion; if $f \leqslant r_i$, the node is not convinced and keeps his/her own opinion.

The resistance parameter r_i can be different or the same for all nodes. One can also play around with the fraction of nodes that is stubborn, for example, there can be a fraction of nodes that is completely stubborn ($r = 1$) and a remaining fraction that is not stubborn at all ($r = 0$). Many other possibilities can be found according to how the stubborn people are distributed and according to how stubborn they are made. This makes the model very flexible.

The threshold model

In the threshold model one introduces a notion of stubbornness by introducing a threshold parameter T. This parameter can be either the same for all the nodes (T) or can be made individual (T_i). The main idea is that nodes can only change their opinion to the other opinion if the fraction of friends with that opinion (p) is bigger than T ($p > T$). The idea behind this threshold model is that people are a bit hesitant to change their opinion, but if they have enough friends with the other opinion, they might get convinced of that opinion as well. The more persistent you are, the more friends you need in order to be convinced of another idea.

There are different possibilities of implementing this kind of model. One way, which makes use of the majority model, is as follows. First the fraction of friends with opinion A and opinion B, p_A and p_B respectively, are determined; then the following rules are used to update the opinion

- If $p_A > p_B$ and $p_A > T$, opinion A is adopted,

- If $p_B > p_A$ and $p_B > T$, opinion B is adopted,

- In all other cases, the current opinion of the node is not changed.

If $T = 0$, this model reduces to the pure majority model described in Section 3.1.1. If $T = 1$, the nodes don't change their opinion. For $0 < T < 1$, the model is a majority model, but the nodes are a bit more persistent towards their own opinion and need a larger fraction of friends to convince them to switch to the other viewpoint.

Note that the above examples are only a handful of the possible ways to include some notion of stubbornness in the model under study, many more exist or can be invented.

Chapter 4

Materials and methods in this thesis

The previous two chapters, Chapter 2 and 3, gave a broad introduction/overview of networks and opinion dynamics models respectively. In this chapter we will summarize the specific networks and opinion dynamics models that are used in this thesis.

Section 4.1 gives an overview of the network models that are used in this thesis. In chapter 2, Section 2.2 we already discussed their main properties; here we will mainly discuss why we use these networks. Which of their properties are of interest in this thesis? What are there benefits and possible shortcomings regarding to our goals?

In section 4.2, we will discuss the real-world networks used in this thesis. The reasons of choosing these specific networks will be explained, together with there possible strengths and weaknesses.

Finally, Section 4.3 discusses the opinion dynamics model that is implemented in this thesis. Again we will give some focus on possible advantages and disadvantages of the model.

4.1 Network models

This thesis uses three main network models, namely

- The Erdős-Rényi model (ER),

- The Watts-Strogatz model (WS),

- The stochastic block model, of which two types are implemented

  - The regular stochastic block model (SBM),
  - The stochastic Watts-Strogatz block model (SBM-WS).

Each of these networks serve specific goals which will shortly be discussed in the following paragraphs.

The Erdős-Rényi model

The Erdős-Rényi model is mainly used as a benchmark model. In this thesis we focus intensively on two network features/topologies that may impact the formation of echo chambers, namely a high clustering coefficient and a community structure. It is therefore necessary to have a model that lacks

both of these features to compare the results with. The Erdős-Rényi model is one of the most simple random network models that doesn't display both of these features and is hence the perfect comparison model.

It is expected that the random nature of the model and the lack of a high clustering and a community structure will result in the absence of echo chambers.

The Erdős-Rényi model is also used in the regular stochastic block model to construct the constituent communities.

The Watts-Strogatz model

The Watts-Strogatz model is used in this thesis as a general model to investigate the effect of high clustering on the evolution of opinions and the formation of echo chambers.

In the paper *"Modelling opinion dynamics in the age of algorithmic personalization"* (see reference [45]) Prof. Perra and Prof. Rocha already showed that a high clustering coefficient enhances the formation of echo chambers [45]. Here, we will use the Watts-Strogatz model to investigate the interplay between high clustering and the presence of stubborn people and, on the other hand, we will use it as a comparison for networks that posses a community structure without having a high clustering coefficient. We will investigate whether a community structure behaves similar to a high clustering regarding the appearance of echo chambers.

Beside the above, the Watts-Strogatz model is also used to build the constituent communities in the stochastic Watts-Strogatz block model.

The stochastic block model

One of the main features of the stochastic block model is the ability to generate networks with a community structure. We will therefore use this model to investigate the impact of community structure/-modularity on the evolution of opinions and the formation of echo chambers.

It is expected that a community structure will enhance the formation of echo chambers compared to the random network case. The question remains, however, whether the influence of a community structure on the formation of echo chambers is of the same caliber as that of a high clustering. It is also of interest to investigate the possible relation between the appearance of echo chambers and the modularity of the network.

The stochastic Watts-Strogatz block model is used as a model that has both a high clustering and a community structure. We will check if the combination of both features displays any difference in the formation of echo chambers compared to having only a single feature (either community structure or high clustering).

All the above models have a homogeneous degree distribution. In this thesis, we didn't focus on network models which generate networks with a heterogeneous degree distribution such as the Barabási-Albert model. It might, however, be an interesting extension to investigate such types of models as well.

The above models obviously have their added value and benefits. One of their most important plus points is their simplicity. They are the most simple network models that one can create and each

of them have characteristics and topologies that can be easily tuned (clustering for WS, community structure for SBM). This makes it really easy to investigate the effect of different network topologies (clustering, community structure) without the need of taking the interplay with other possible complex network features and topologies into account.

This plus side is, however, also one of the possible downsides. Real networks aren't simple networks with single, isolated characteristics, but are instead complex networks where many different characteristics and topologies may each play a role. In other words, network models may lack many features of real-world networks. It is thus always a bit tricky, and sometimes even incorrect, to extrapolate the results that are obtained by using network models to real-world situations.

The fact that all the above models have a homogeneous degree distribution may also be a drawback, since most real-world networks have a heterogeneous degree distribution.

The conclusion of this section is that all the network models used in this thesis serve a goal and have specific topologies that are of interest in this thesis. But, it must always be kept in mind that, as the name itself already says, these models are just models of the real world. Results that are obtained from them should not be thoughtlessly extrapolated to real-world networks.

4.2 Real-world networks

Ideally, we would test the opinion dynamics model that was used throughout this thesis on network data directly obtained from social network companies. However, most of these companies are not to keen on sharing any form of data. This makes it very hard to find accurate and complete social media network data. Hence, we chose to test the model on other types of social networks, either on-line or off-line. Some possibilities are any form of friendship networks, e-mail networks, collaboration networks, etc.

One of the decisive properties of the real-world networks that were eventually used, was there size. The main, and actually only, reason for this was the computation time. If the networks would be to large, it would take weeks, sometimes even months before there would be any kind of results.

The size of the network models that were used, was around a thousand nodes. So for the real-world networks, we also used networks with a size of magnitude of thousand. This obviously benefits the computation time; it is, however, unknown how the network size affects the results. It is therefore recommended to test the model on larger network models and on larger real-world networks in future researches on this topic.

4.3 Opinion dynamics models

Chapter 5

Results

In this chapter the results are shared and discussed. The chapter in subdivided into four main sections

- The impact of network structure: the effect of different network topologies on the evolution of the opinions and the formation of echo chambers is reviewed. Main focus lies on networks with a high clustering coefficient and/or a community structure.

- The impact of stubborn people: stubborn people are introduced in the system; in this section the effects of those stubborn people on the evolution of opinions and the formation of echo chambers will be discussed.

- Random versus community distributed opinions/stubborn agents: this section investigates different ways of distributing the opinions and/or the stubborn agents in a network model with a community structure (SBM).

- Real-world networks: the opinion dynamics model of this thesis will be ran on some real-world networks. This section will discuss the results and compare them to results obtained from network models with a similar average degree, clustering coefficient and/or community structure.

5.1   Impact of network structure

In this section we will discuss some results related to the network structure and network topologies of the different network models that are used in this thesis. We will mainly focus on the effect of a high clustering coefficient and a community structure. Two different opinion distributions are investigated

- An initial $50/50$ opinion distribution: the initial fraction of nodes with opinion 0 is equal to $P_0(0) = 0.5$ and the initial fraction of nodes with opinion 1 is equal to $P_1(0) = 0.5$.

- An initial $20/80$ opinion distribution: the initial fraction of nodes with opinion 0 is equal to $P_0(0) = 0.2$ and the initial fraction of nodes with opinion 1 is equal to $P_1(0) = 0.8$. Hence, we are dealing with a minority and a majority opinion.

For both opinion distributions we investigate the evolution of the opinion fractions with time and the formation of echo chambers.

Before sharing the results, a quick disclaimer: Some of the results that are shared in this section, were already obtained by prof. Nicola Perra and prof. Luis E.C. Rocha in their paper *"Modelling opinion dynamics in the age of algorithmic personalization"* (see reference [45]). These results include the effect of a high clustering coefficient (Watts-Strogatz model) and a random network model. Similar results (albeit with a different network size and average degree) are included in this section to give a complete overview and to be able to compare a community structure to a high clustering coefficient.

5.1.1 Starting conditions $P_0(0) = 0.5$ and $P_1(0) = 0.5$

We start with the following opinion distribution: a fraction of nodes with opinion 0 equal to 0.5 ($P_0(0) = 0.5$) and a fraction of nodes with opinion 1 equal to 0.5 ($P_1(0) = 0.5$). The opinions are randomly distributed over the network. The opinion dynamics model discussed in Chapter 4, Section 4.3 is used. We set the activation probability equal to $p_{act} = 0.1$ and the actual time-line of each user has a maximum length equal to $S = 20$.

Four different network models are implemented, each with a total number of nodes equal to $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$

- The Erdős-Rényi model (ER) with edge probability $p = 0.01$.

- The Watts-Strogatz model (WS) with mean degree $K = 10$; two different values of the rewire probability β are used, each of them has a slightly different value of the average clustering coefficient, namely $\beta = 0.01$ with $\langle C \rangle \sim 0.60$ and $\beta = 0.06$ with $\langle C \rangle \sim 0.55$. This model is used to investigate the effect of clustering.

- The stochastic block model with 10 communities of size 100; two networks with different modularities Q are investigated: a high modularity network, $Q \sim 0.9$, with an edge probability inside the communities equal to $p_{cl} = 0.1$ and an edge probability between communities equal to $p_{add} = 0.001$, and a low modularity network ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$). This model is used to investigate the effect of a community structure.

- The stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100. Each community has the following Watts-Strogatz parameters: mean degree $K = 10$ and rewire probability $\beta = 0.01$. Edges are added between communities with a probability $p_{add} = 0.001$. This is a network with both a high clustering ($\langle C \rangle \sim 0.55$) and a high modularity ($Q \sim 0.9$).

Evolution of the opinion fractions

Figure 5.1 shows the results for the evolution of the fraction of opinion 0 over time, together with the standard deviation σ. Two different filtering algorithms are compared to each other, namely the PR method and the REC method (see Chapter 3, Section 3.3 for a detailed explanation). None of the network topologies, nor the filtering algorithms, is able to break the initial $50/50$ opinion distribution. Having a high clustering or a clear community structure has no impact on the prevalence of the opinions. The status quo is maintained and the system is not pushed towards a different equilibrium.
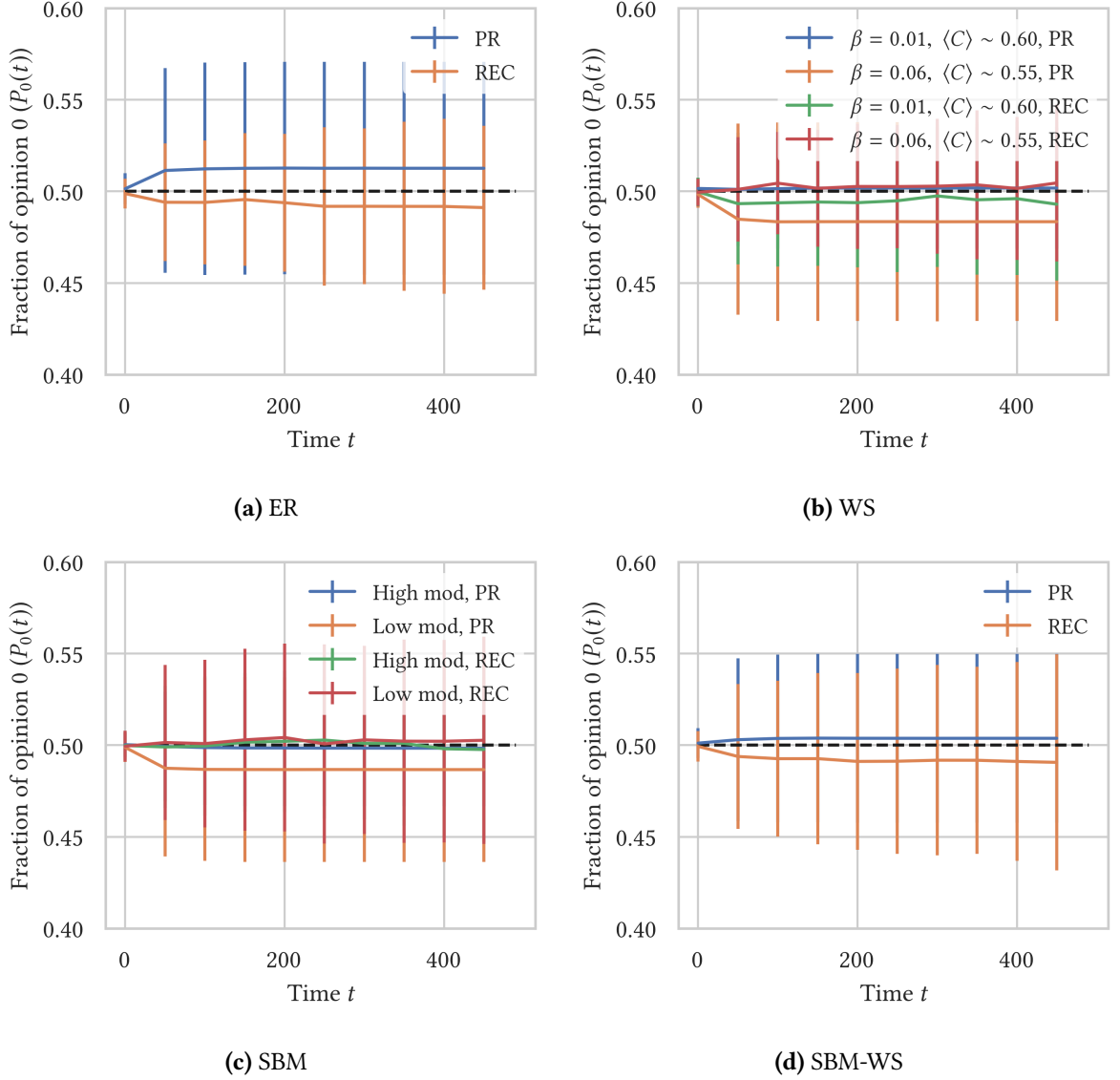
(a) ER

(b) WS

(c) SBM

(d) SBM-WS

Figure 5.1: Evolution of the fraction of nodes with opinion 0 ($P_0(t)$) over time for an initial 50/50 distribution. Four different network models are used (each with an average degree around $\langle k \rangle \sim 10$ and a total number of nodes $N = 10^3$): **(a)** Erdős-Rényi model (ER), $p = 0.01$; **(b)** Watts-Strogatz model (WS), $K = 10$, two slightly different values of the clustering coefficient are used ($\langle C \rangle \sim 0.55$ and $\langle C \rangle \sim 0.60$); **(c)** Stochastic block model (SBM) with 10 communities of size 100, a high modularity case ($p_{\text{cl}} = 0.1$, $p_{\text{add}} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{\text{cl}} = 0.03$, $p_{\text{add}} = 0.008$, $Q \sim 0.2$) are depicted; **(d)** Stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{\text{add}} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$. For each model two different filtering algorithms are compared to each other: the REC and the PR filtering algorithm. Each plot is the average of 10^2 independent simulations and data points are shown for every 50 time steps to improve visualization.

Echo chambers

(a) REC   **(b)** PR
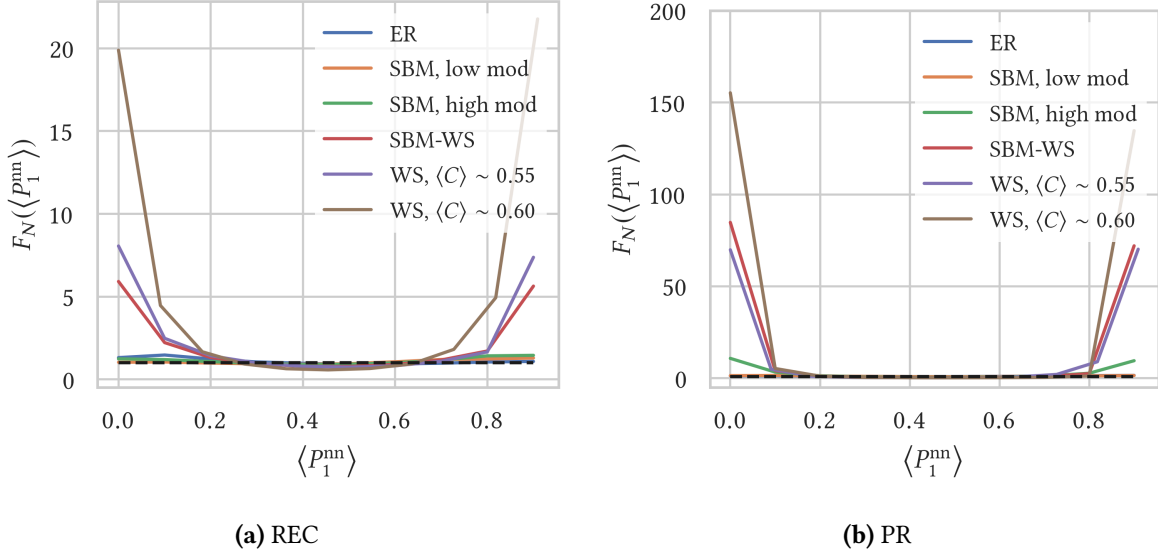
Figure 5.2: Distribution of the fraction of friends (nearest neighbors, nn) of node i with the same opinion 1 at $t = 500$. The distribution is normalized by dividing for the same quantity computed at $t = 0$. **(a)** Results for the REC method; **(b)** Results for the PR method. Four different network models are used, each with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$: The Erdős-Rényi model (ER) with $p = 0.01$; The stochastic block model (SBM) with 10 communities of size 100, a high modularity case ($p_{cl} = 0.1$, $p_{add} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$) are depicted; The stochastic Watts-Strogatz model (SBM-WS) with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{add} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$; The Watts-Strogatz model (WS) with $K = 10$, two slightly different values of the clustering coefficient are used ($\langle C \rangle \sim 0.55$ and $\langle C \rangle \sim 0.60$). The starting conditions are $P_0(0) = 0.5$ and $P_1(0) = 0.5$. Each plot is the average of 10^2 independent simulations.

Figure 5.2 shows the distribution of the fraction of friends (nearest neighbors, nn) of node i with the same opinion 1 at $t = 500$, $\langle P_1^{nn} \rangle$. The value shown on the y-axis, $F_N(\langle P_1^{nn} \rangle)$, is the result of dividing the average distribution of opinions in each neighborhood at $t = 500$ by its corresponding value at $t = 0$. The region on the x-axis close to zero describes neighborhoods in which none, or very few users, adopt opinion 1 (these are thus neighborhoods in which most users adopt opinion 0). Conversely, the region on the x-axis close to one describes neighborhoods in which a majority of the users adopt opinion 1. The distribution is obtained by imposing ten equispaced bins, $\Delta x = 0.1$ [45]. Error bars are omitted for the sake of an improved visualization.

These types of plots are perfectly suited to investigate the formation of opinion clusters (echo chambers) and effects such as polarization in the system under study.

In Figure 5.2 two different filtering algorithms are compared, the REC and the PR method. The results for both filtering algorithms look qualitatively the same. The main difference is that the formation of echo chambers, using the PR method, is about ten times as big compared to the formation of echo chambers using the REC method.

The Erdős-Rényi model (ER) showns no, or a very limited, sign of echo chambers, for both the REC

method and the PR method. Networks with a high clustering (WS and SBM-WS), on the other hand, show a clear sign of polarization and formation of echo chambers (visible in the two spikes, one around zero and one around one). Going to slightly higher values of the clustering coefficient ($\langle C \rangle \sim 0.55$ to $\langle C \rangle \sim 0.60$) in the Watts-Strogatz model results in significantly larger echo chambers. This suggests that the formation of echo chambers is really sensitive to changes in the clustering coefficient. The stochastic block model (SBM) with a high modularity, hence, a clear community structure, does show an appearance of polarization (mainly when using the PR filtering algorithm), however, the formation of echo chambers is significantly smaller than compared to networks with a high clustering. The stochastic block model with a low modularity has no, or very limited, echo chambers, similar as for the Erdős-Rényi model. This is to be expected, since a stochastic block model with a low modularity, and hence a very blurred, almost non-existing community structure, strongly resembles an Erdős-Rényi network. The stochastic Watts-Strogatz block model (SBM-WS) has both a community structure and a high average clustering. The formation of echo chambers when using this model is approximately of the same size compared to the formation of echo chambers in a Watts-Strogatz model with a similar average clustering. This suggests that it is mainly the high clustering that drives the system to form echo chambers, while the community structure plays a less important role. However, as said before, the stochastic block model with a high modularity also has a (smaller) formation of echo chambers, so having a clear and strong community structure definitely plays some role in phenomena such as polarization and the formation of opinion clusters.

Note that for an initial $50/50$ opinion distribution none of the network models, nor the filtering algorithms could break this status quo (see Figure 5.1). Despite this preserving of the balanced opinion distribution, the interplay of filtering algorithms and network structure can potentially shift the initial randomly distributed opinions to a more clustered opinion distribution and thus lead to polarization and formation of echo chambers.

5.1.2 Starting conditions $P_0(0) = 0.2$ and $P_1(0) = 0.8$

The initial fraction of nodes with opinion 0 is now equal to $P_0(0) = 0.2$, opinion 0 thus becomes the minority opinion. The initial fraction of nodes with opinion 1 is equal to $P_1(0) = 0.8$, hence this is the majority opinion. The opinions are, initially, randomly distributed over the system. Again, the opinion dynamics model of Chapter 4, Section 4.3 is implemented; the activation probability is set equal to $p_{\text{act}} = 0.1$ and the actual time-line of each user has a maximum length $S = 20$.

Once again, we look at four different network models each with a total number of nodes equal to $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$

- The Erdős-Rényi model (ER) with edge probability $p = 0.01$.

- The Watts-Strogatz model (WS) with mean degree $K = 10$ and rewire probability $\beta = 0.06$. The value of the average clustering coefficient is equal to $\langle C \rangle \sim 0.55$. This model is used to investigate the effect of clustering.

- The stochastic block model with 10 communities of size 100; two networks with different modularities Q are investigated: a high modularity network ($p_{cl} = 0.1$, $p_{add} = 0.001$, $Q \sim 0.9$) and a low modularity network ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$). This model is used to investigate the effect of a community structure.

- The stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100. Each community has the following Watts-Strogatz parameters: mean degree $K = 10$ and rewire probability $\beta = 0.01$. Edges are added between communities with a probability $p_{add} = 0.001$. This is a network with both a high clustering ($\langle C \rangle \sim 0.55$) and a high modularity ($Q \sim 0.9$).

Evolution of the opinion fractions

Figure 5.3 shows the results for the evolution of the fraction of nodes with opinion 0, which is the minority opinion in this case, over time, together with the standard deviation σ. For each network model two filtering algorithms are compared, namely the REC filtering algorithm and the PR filtering algorithm.

The REC method is not able to break the prevalence of opinion 0 over time for neither of the four different network models. When using the PR method, the results look very different. The prevalence of the minority opinion, $P_0(t)$, decreases over time. This decrease is a direct consequence of the PR filtering algorithm. Since the PR method orders the posts according to the users' preferences, nodes with the majority opinion 1 will mainly get to see posts with the similar opinion 1 on their time-line, which limits the chance of changing their opinion to the minority opinion. Nodes with the minority opinion 0, on the other hand, will obviously see posts with opinion 0 on their time-line when using the PR method. However, since there are more nodes with opinion 1 than nodes with opinion 0, the nodes with the minority opinion 0 will also get to see a significant number of posts with the majority opinion 1 on their time-line. The conclusion is that, when using the PR method, there is an imbalance in which opinion can switch to which: nodes with the minority opinion 0 have a chance of switching to the majority opinion 1, whereas nodes with the majority opinion 1 have (almost) no chance of switching to the minority opinion 0. This imbalance leads to the decrease of the minority opinion 0.

From Figure 5.3, it is clear that the minority opinion 0 eventually ceases to exist for all the network models, $P_0(500) \sim 0$. This result differs from the conclusion that prof. N. Perra and prof. L.E.C. Rocha made in the paper *"Modelling opinion dynamics in the age of algorithmic personalization"* (see reference [45]). They found that the minority opinion was able to survive over time, even for the PR method, and reported a minority opinion fraction at $t = 500$ between $P_0(500) = 0.06$ and $P_0(500) = 0.09$ [45]. However, the results that they obtained are for networks with a total number of nodes equal to $N = 10^5$ and an average degree around $\langle k \rangle \sim 6$. In this thesis, on the other hand, results are obtained for $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$. In both cases the maximum length of the actual time-line is set equal to $S = 20$. The difference in the results is most likely an interplay between the network size N, the average degree $\langle k \rangle$ and the maximum length of the actual time-line S. This can for example be seen in Figure 5.4, which shows the results for the evolution of the fraction of opinion 0 for an initial 20/80 opinion distribution and using the PR method; Figure 5.4a displays the results for a WS model with $N = 10^3$, $K = 6$ and $\beta = 0.01$; Figure 5.4b shows the results for a WS model with $N = 10^4$, $K = 6$ and $\beta = 0.01$. In both cases the average clustering coefficient is around $\langle C \rangle \sim 0.58$

and the maximum length of the actual time-line is $S = 20$. In this figure we can see that the minority opinion 0 doesn't vanish from the system over time, but instead reaches a value of $P_0(500) \sim 0.027$ in Figure 5.4a and a value of $P_0(500) \sim 0.03$ in Figure 5.4b.
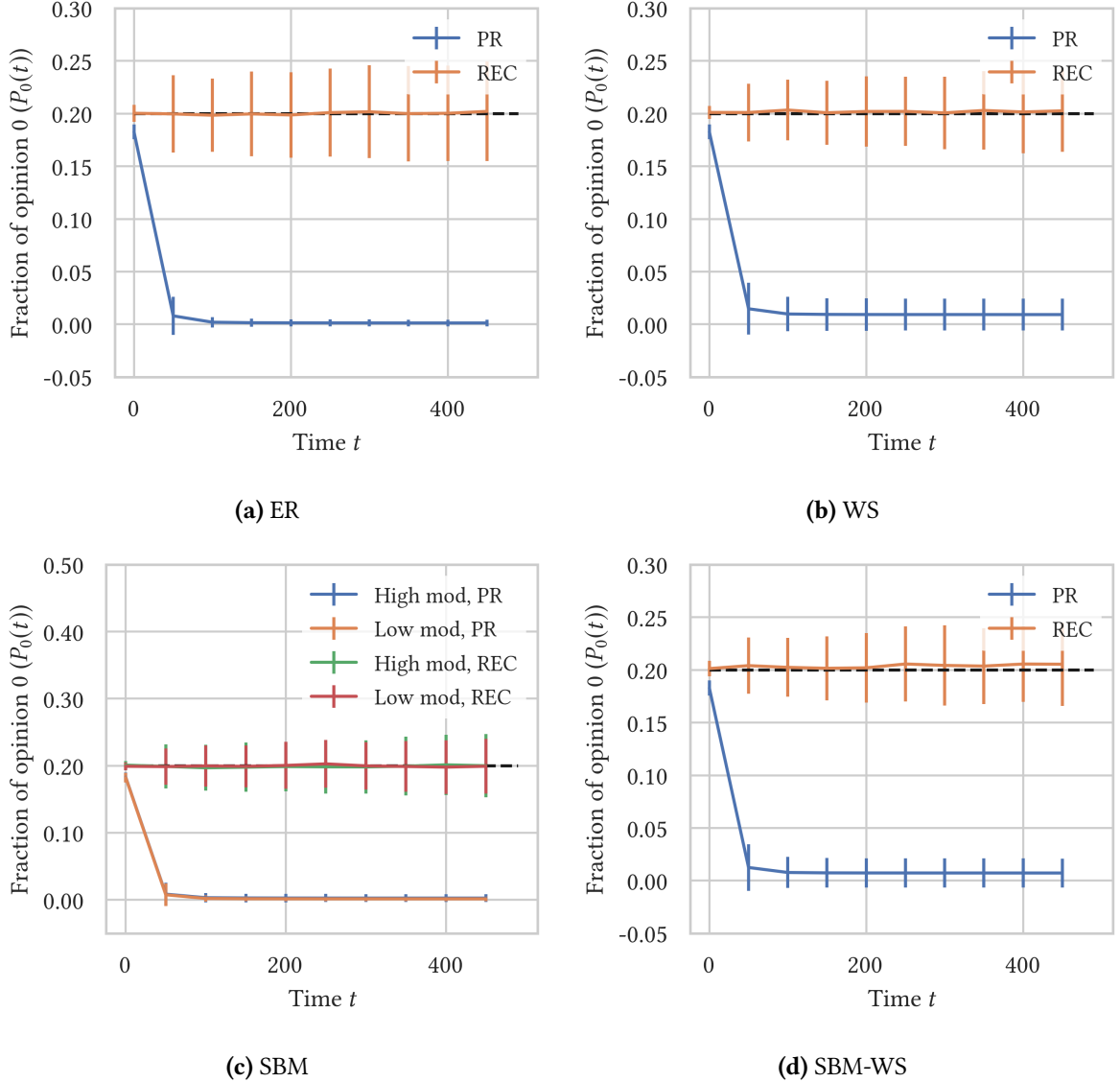

(a) ER

(b) WS

(c) SBM

(d) SBM-WS

Figure 5.3: Evolution of the fraction of nodes with opinion 0 ($P_0(t)$) over time for an initial $20/80$ distribution. Four different network models are used (each with an average degree around $\langle k \rangle \sim 10$ and a total number of nodes $N = 10^3$): **(a)** Erdős-Rényi model (ER), $p = 0.01$; **(b)** Watts-Strogatz model (WS) with $K = 10$, $\beta = 0.06$ and $\langle C \rangle \sim 0.55$; **(c)** Stochastic block model (SBM) with 10 communities of size 100, a high modularity case ($p_{cl} = 0.1$, $p_{add} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$) are depicted; **(d)** Stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{add} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$. For each model two different filtering algorithms are compared to each other: the REC and the PR filtering algorithm. Each plot is the average of 10^2 independent simulations and data points are shown for every 50 time steps to improve visualization.
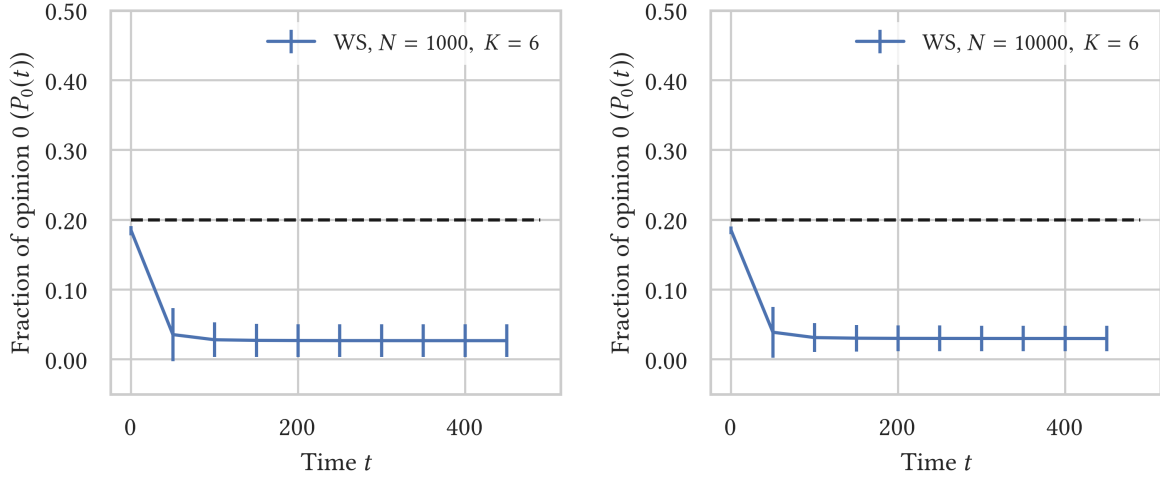
(a) WS, $N = 10^3$, $K = 6$, $\beta = 0.01$ **(b)** WS, $N = 10^4$, $K = 6$, $\beta = 0.01$

Figure 5.4: Evolution of the fraction of nodes with opinion 0 ($P_0(t)$) over time for an initial 20/80 distribution. The Watts-Strogatz network is depicted with **(a)** $N = 10^3$, $K = 6$, $\beta = 0.01$ and **(b)** $N = 10^4$, $K = 6$, $\beta = 0.01$. In both cases the average clustering coefficient is around $\langle C \rangle \sim 0.58$ and the PR filtering method is used. Each plot is the average of 10^2 independent simulations and data points are shown for every 50 time steps to improve visualization.

Echo chambers

Figure 5.5 shows the distribution of the fraction of friends of node i with the same opinion 1 at $t = 500$ for the four network models: Erdős-Rényi model (ER), Watts-Strogatz model (WS), stochastic block model (SBM) and stochastic Watts-Strogatz block model (SBM-WS). For each network, the REC and PR method are compared to each other. The interpretation of the figure is the same as was described in Subsection 5.1.1, the echo chamber paragraph.

The results for the Erdős-Rényi model and the stochastic block model (both the high and low modularity) look qualitatively the same. Similarly, the results for the Watts-Strogatz model and the stochastic Watts-Strogatz model agree with each other as well. This difference in results between ER/SBM (negligible clustering) and WS/SBM-WS (high clustering) suggests that it is mainly the clustering that drives these differences.

There is no visible effect for having a clear community structure. This can be clearly seen in Figure 5.5c where the results for a stochastic block model with a high modularity (clear community structure) and a low modularity (negligible community structure) are compared. There is (almost) no difference in the results for a high and a low modularity.

Lets first discuss the results for the Erdős-Rényi model and the stochastic block model. When using the REC method, the formation of echo chambers around the majority opinion 1 is hampered. This is somewhat surprising, since 80% of the nodes carry this opinion and is probably a consequence of the absence of clustering and the presence of short average path lengths in these models. There is even a little spike around $\langle P_1^{nn} \rangle = 0.1$, which indicates an increase in the fraction of nodes with the majority of their neighbors carrying the minority opinion 0. From figure 5.3 it was already concluded that the REC method doesn't break the prevalence of the opinions. The opinion distribution is however shifted from the case where the minority opinion 0 is more or less randomly distributed over the network to the case where the minority opinion 0 groups together in opinion clusters.

The PR method shows a completely different tendency. There is an increase in the fraction of nodes with (almost) all their neighbors carrying the majority opinion 1. This is a consequence from the fact that the PR method reduces the visibility of the minority opinion 0 in the neighborhood of nodes with the majority opinion 1.

Networks with a high clustering coefficient (WS and SBM-WS) behave completely different. For both the REC and PR method the high clustering enhances the formation of echo chambers around the minority opinion 0 (clearly visible in the spike around $\langle P_1^{nn} \rangle = 0$ in both Figure 5.8b and Figure 5.5d).
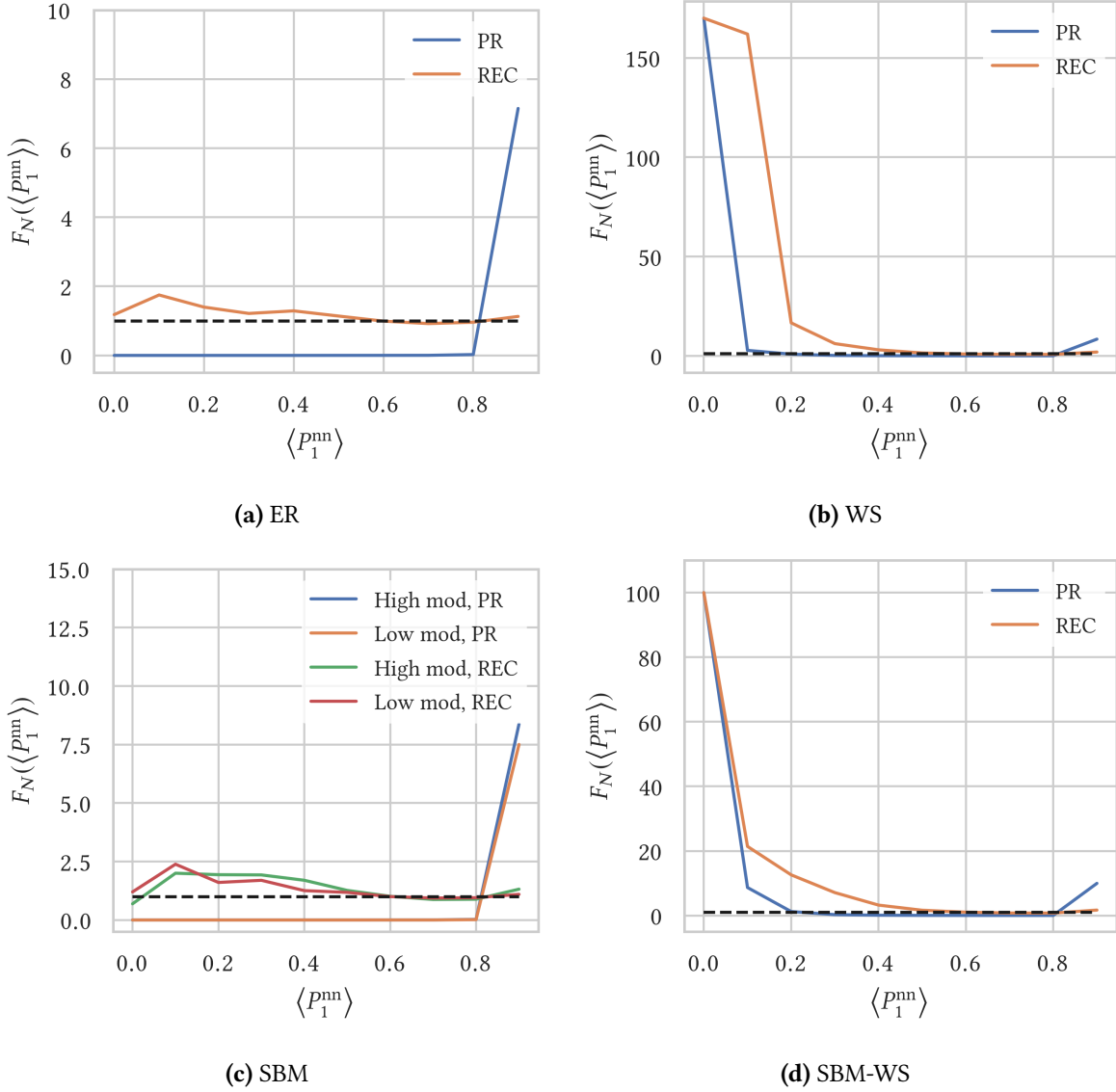
Figure 5.5: Distribution of the fraction of friends (nearest neighbors, nn) of node i with the same opinion 1 at $t = 500$. The distribution is normalized by dividing for the same quantity computed at $t = 0$. Four different network models are used, each with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$: **(a)** Erdős-Rényi model (ER), $p = 0.01$; **(b)** Watts-Strogatz model (WS) with $K = 10$, $\beta = 0.06$ and $\langle C \rangle \sim 0.55$; **(c)** Stochastic block model (SBM) with 10 communities of size 100, a high modularity case ($p_{\text{cl}} = 0.1$, $p_{\text{add}} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{\text{cl}} = 0.03$, $p_{\text{add}} = 0.008$, $Q \sim 0.2$) are depicted; **(d)** Stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{\text{add}} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$. For each model two different filtering algorithms are compared to each other: the REC and the PR filtering algorithm. The starting conditions are $P_0(0) = 0.2$ and $P_1(0) = 0.8$. Each plot is the average of 10^2 independent simulations.

5.1.3 Relation between formation of echo chambers and community structure
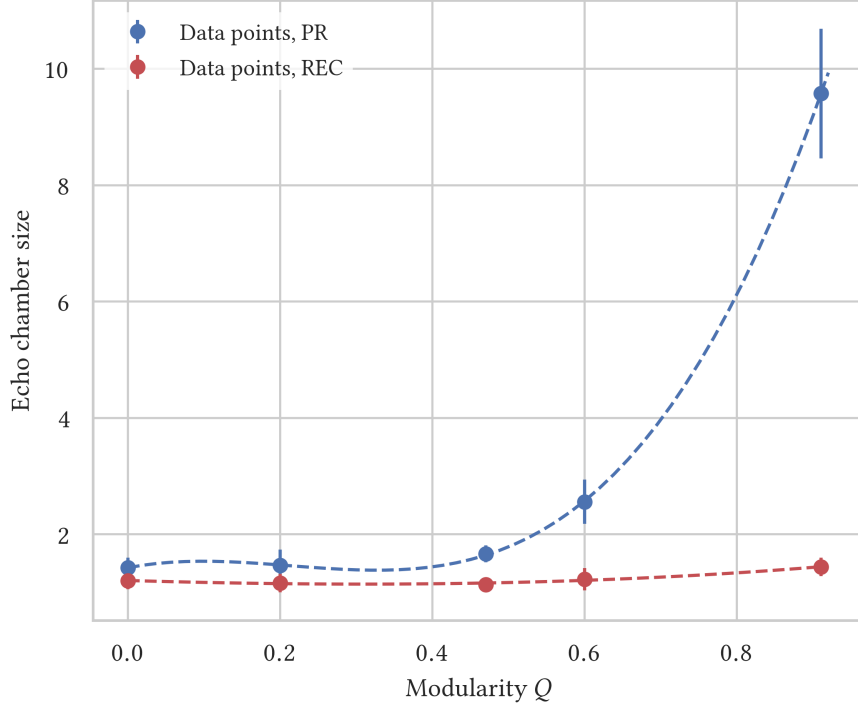

Figure 5.6: The relation between the formation of echo chambers and the community structure, represented by the modularity Q, for both the PR and REC filtering algorithms. The y-axis shows the average of the normalized fraction of nodes with all neighbors having opinion 0 and the normalized fraction of nodes with all neighbors having opinion 1. This value represents the formation of echo chambers in the system (denoted with echo chamber size in the figure). The initial opinion distribution is $50/50$. The data point with $Q = 0$ is the result for an ER network with $N = 10^3$, $p = 0.01$ and an average degree around 10. The other four data points are the results for SBM's with 10 communities of size 100 (total number of nodes $N = 10^3$), each with a different modularity (the modularity is changed by changing the edge probabilities) and an average degree around $\langle k \rangle \sim 10$. Each data point is the result of 10^2 independent simulations.

Now, lets look a bit closer to the relation between the formation of echo chambers and the community structure, which is quantified by the modularity Q. In Figure 5.6, the size of the echo chambers versus the modularity is depicted. The y-axis represents the average of the normalized fraction of nodes with all neighbors carrying opinion 0 and the normalized fraction of nodes with all neighbors carrying opinion 1 (with normalized we mean that the computed value for the fraction of nodes with all neighbors having opinion 0 or all neighbors having opinion 1 at $t = 500$ is divided by the corresponding value at $t = 0$). This is a measure for the formation of echo chambers (opinion clusters) in the system. The data points represent the value of the echo chamber size in networks with a different modularity. The initial opinion distribution was always $50/50$.

In increasing order of modularity, the used network models are (all with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$)

- $Q = 0 \rightarrow$ Erdős-Rényi model with edge probability $p = 0.01$,

- $Q \sim 0.2 \rightarrow$ SBM with 10 communities of size 100; the edge probability inside the communities is $p_{\text{cl}} = 0.03$ and the edge probability between communities is $p_{\text{add}} = 0.008$,

- $Q \sim 0.47 \rightarrow$ SBM with 10 communities of size 100; the edge probability inside the communities is $p_{\text{cl}} = 0.05$ and the edge probability between communities is $p_{\text{add}} = 0.005$,

- $Q \sim 0.6 \rightarrow$ SBM with 10 communities of size 100; the edge probability inside the communities is $p_{\text{cl}} = 0.07$ and the edge probability between communities is $p_{\text{add}} = 0.004$,

- $Q \sim 0.9 \rightarrow$ SBM with 10 communities of size 100; the edge probability inside the communities is $p_{\text{cl}} = 0.1$ and the edge probability between communities is $p_{\text{add}} = 0.001$.

The data points are shown together with the standard error $SE = \sigma/\sqrt{N}$. The standard error SE is chosen here instead of the standard deviation σ. This is because the standard deviation is quite large. Each data point is the average of 10^2 independent simulations. The dispersion/variation, which is measured by the standard deviation, in the set of the 10^2 obtained values is quite large. However, if we run the 10^2 simulations over and over again, each time calculating the average, the obtained value of the average is always more or less the same. This suggests that the value of the calculated average lies close to the real average. Since we are mostly interested in the correctness of the value of the average, which is quantified by the standard error, we chose to depict the results together with this standard error.

The PR filtering algorithm is compared to the REC filtering algorithm. For most values of the modularity, the REC method produces almost no echo chambers. If the modularity becomes high enough we start to see a very small increase in the formation of echo chambers, however even for $Q \sim 0.9$, the size of the echo chambers remains smaller than 2, which means that after 500 time steps there are less than twice as much echo chambers than in the beginning at $t = 0$.

For the PR method, the increase in the size of the echo chambers with increasing modularity is much more significant. From a modularity approximately equal to $Q \sim 0.5$, we can see a sharp increase in the formation of echo chambers. For modularity values smaller than $Q \sim 0.5$, the formation of echo chambers remains very small, even when using the PR method.

These results strongly suggest that a community structure enhances the formation of echo chambers in the system, mostly when using the PR filtering algorithm. A quite clear community structure, represented by a high modularity, is however needed to really see the effect of this community structure on the formation of opinion clusters.

5.2   Impact of stubborn people

In this section, the impact of introducing stubborn people will be reviewed. We focus on two different models, both are thoroughly discussed in Chapter 3, Section 3.5.2

- Including a stubbornness parameter: these results are discussed in Section 5.2.1,

- The majority threshold model: results for this model are given in Section 5.2.2.

5.2.1   Including a stubbornness parameter

Here, we review the results of introducing stubborn people by including a stubbornness parameter r_i. We look at two slightly different case

- Making a fraction of the nodes (going from none of the nodes to all the nodes) completely stubborn ($r_i = 1$),

- Making all nodes partly stubborn (going from not stubborn at all, $r_i = 0$ to completely stubborn, $r_i = 1$).

For both cases we will look at both the evolution of the opinions and the formation of echo chambers. We will again implement four different network models in order to investigate the possible interplay between network structure and stubbornness, these models are (each with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$)

- The Erdős-Rényi model (ER) with edge probability $p = 0.01$.

- The Watts-Strogatz model (WS) with mean degree $K = 10$ and rewire probability $\beta = 0.06$, with $\langle C \rangle \sim 0.55$. This model is used to investigate the effect of clustering.

- The stochastic block model with 10 communities of size 100; two networks with different modularities Q are investigated: a high modularity network ($p_{cl} = 0.1$, $p_{add} = 0.001$, $Q \sim 0.9$) and a low modularity network ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$). This model is used to investigate the effect of a community structure.

- The stochastic Watts-Strogatz block model (SBM-WS) with 10 communities of size 100. Each community has the following Watts-Strogatz parameters: mean degree $K = 10$ and rewire probability $\beta = 0.01$. Edges are added between communities with a probability $p_{add} = 0.001$. This is a network with both a high clustering ($\langle C \rangle \sim 0.55$) and a high modularity ($Q \sim 0.9$).

For each network model, we compare the REF, REC and PR method to each other to analyze the possible interplay between stubbornness and filtering algorithm.

Making a fraction of the nodes completely stubborn

We start by making a fraction of the nodes (fracRes) completely stubborn ($r_i = 1$). We go from no stubborn nodes, fracRes $= 0$, to all the nodes are stubborn, fracRes $= 1$. In all cases, the stubborn nodes are randomly distributed over the network.

Figure 5.7 shows the formation of echo chambers in the system versus the fraction of completely stubborn nodes, fracRes, when using the PR filtering algorithm. The y-axis represents the formation of echo chambers in the same way as was explained for Figure 5.6 in Section 5.1.3. Data points are again shown together with their standard error, SE, for the same reasons as explained in Section 5.1.3. The Erdős-Rényi model (ER) shows no appearance of echo chambers. Introducing a fraction of completely stubborn nodes doesn't change this behavior. The stochastic block model with a low modularity behaves similar as the Erdős-Rényi model (no formation of echo chambers). Note that the data points for the SBM with a low modularity are not visible in Figure 5.7, since they lie behind the ER data points. The other three networks (WS, SBM-WS and SBM with a high modularity) all show qualitatively the same behavior. If an increasing fraction of completely stubborn nodes is introduced, the formation of echo chambers decreases with this increasing fraction of stubborn people. This can be explained as follows. Having a fraction of completely stubborn nodes, reduces the effective number of neighbors that can change their opinion. This in turn leads to the hampering of the formation of echo chambers. The WS and SBM-WS networks are the high clustering networks and for a small fraction of stubborn people they show the largest echo chambers. The stochastic block model with a high modularity has smaller echo chambers than the high clustering networks, but still has some formation of echo chambers for a fraction of stubborn nodes smaller than fracRes $\leqslant 0.1$. This leads us to make the same conclusion as before: a community structure enhances the formation of echo chambers compared to a random network, however not as significant as networks with a high clustering coefficient.
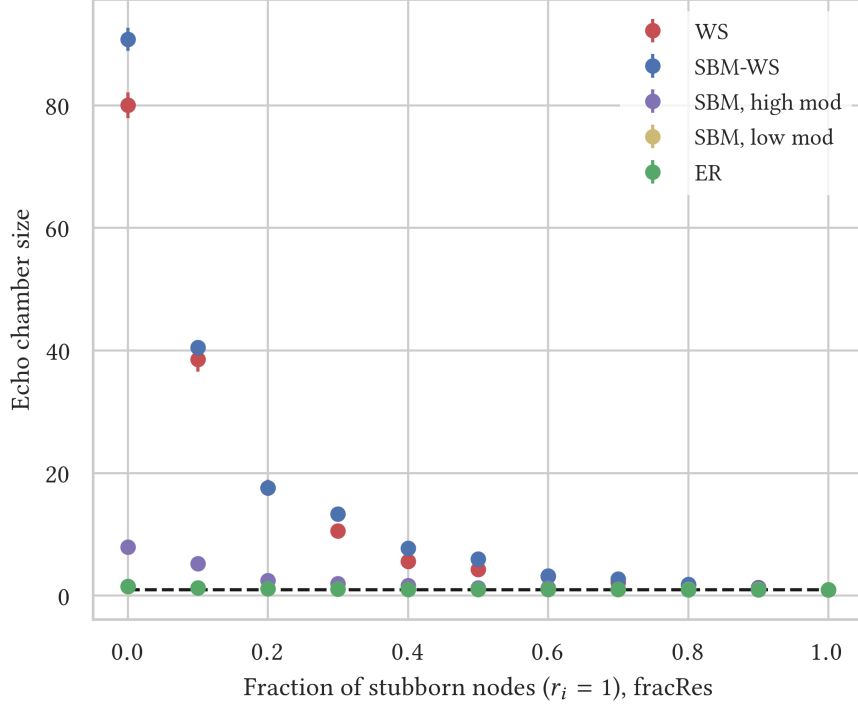
Figure 5.7: Formation of echo chambers versus the fraction of completely stubborn nodes, fracRes, for an initial $50/50$ opinion distribution and using the PR filtering algorithm. The y-axis shows the average of the normalized fraction of nodes with all neighbors having opinion 0 and the normalized fraction of nodes with all neighbors having opinion 1. This value represents the formation of echo chambers in the system (denoted with echo chamber size in the figure). Results are shown for four different network models, each with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$: WS network with $K = 10$, $\beta = 0.06$ and $\langle C \rangle \sim 0.55$; SBM-WS network with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{\text{add}} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$; SBM network with 10 communities of size 100, a high modularity case ($p_{\text{cl}} = 0.1$, $p_{\text{add}} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{\text{cl}} = 0.03$, $p_{\text{add}} = 0.008$, $Q \sim 0.2$) are depicted; ER network with $p = 0.01$. Each data point is the result of 10^2 independent simulations.
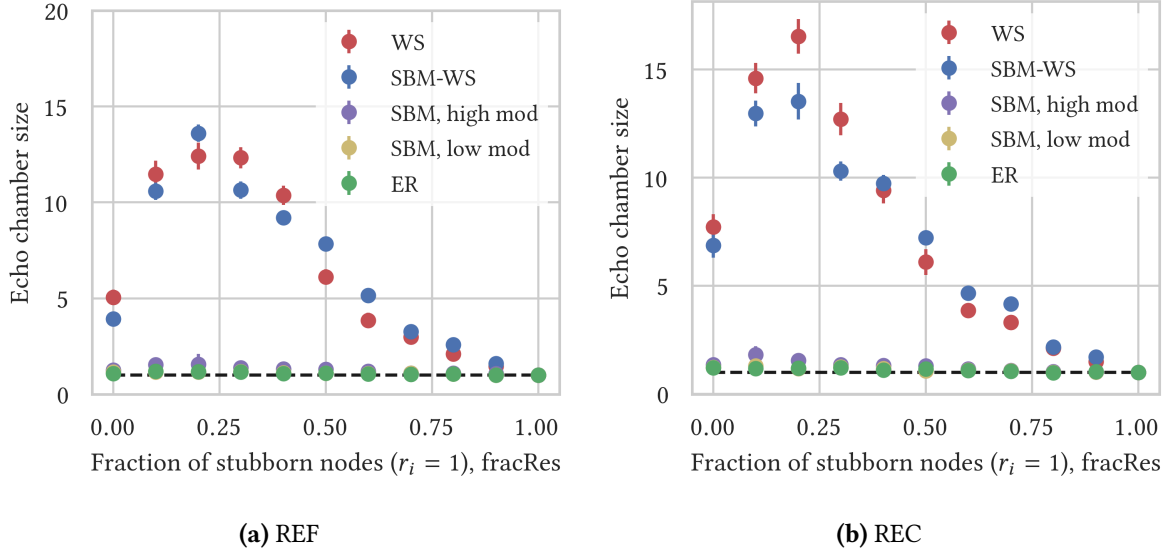
(a) REF

(b) REC

Figure 5.8: Formation of echo chambers versus the fraction of completely stubborn nodes, fracRes, for an initial $50/50$ opinion distribution. **(a)** Using the REF filtering algorithm; **(b)** Using the REC filtering algorithm. The y-axis shows the average of the normalized fraction of nodes with all neighbors having opinion 0 and the normalized fraction of nodes with all neighbors having opinion 1. This value represents the formation of echo chambers in the system (denoted with echo chamber size in the figure). Results are shown for four different network models, each with a total number of nodes $N = 10^3$ and an average degree around $\langle k \rangle \sim 10$: WS network with $K = 10$, $\beta = 0.06$ and $\langle C \rangle \sim 0.55$; SBM-WS network with 10 communities of size 100, each community has the following WS parameters: $K = 10$, $\beta = 0.01$ and $p_{add} = 0.001$, $\langle C \rangle \sim 0.55$ and $Q \sim 0.9$; SBM network with 10 communities of size 100, a high modularity case ($p_{cl} = 0.1$, $p_{add} = 0.001$, $Q \sim 0.9$) and a low modularity case ($p_{cl} = 0.03$, $p_{add} = 0.008$, $Q \sim 0.2$) are depicted; ER network with $p = 0.01$. Each data point is the result of 10^2 independent simulations.

Making all nodes partly stubborn

jsqnbd

5.2.2   The majority threshold model

5.3   Random versus community distributed opinions/stubborn agents

5.4   Real-world networks

Chapter 6

Discussion

Chapter 7

Conclusions and outlook

Appendix A

Network measurements: Derivations

A.1 Modularity

A.1.1 Simple form of network modularity

Using [5]

$$Q_c = \frac{1}{2L} \sum_{(i,j)\in C_c} (A_{ij} - p_{ij}) , \tag{A.1}$$

and [5]

$$p_{ij} = \frac{k_i k_j}{2L} , \tag{A.2}$$

we can write the modularity of the whole network as [5]

$$Q = \frac{1}{2L} \sum_{i,j=1}^{N} \left(A_{ij} - \frac{k_i k_j}{2L} \right) \delta_{c_i,c_j} , \tag{A.3}$$

where c_i represents the community to which node i belongs and where the Dirac-delta function δ_{c_i,c_j} expresses the fact that only pairs of nodes that belong to the same community are taken into account. We can then rewrite the first term as a sum over communities [5]

$$\frac{1}{2L} \sum_{i,j=1}^{N} A_{ij} \delta_{c_i,c_j} = \sum_{c=1}^{n_c} \frac{1}{2L} \sum_{i,j\in C_c} A_{ij} = \sum_{c=1}^{n_c} \frac{L_c}{L} , \tag{A.4}$$

where L_c is the number of edges in community C_c and the factor 2 disappears because each edge is counted twice in A_{ij}. Similarly, the second term in Eq. (A.3) can be rewritten as [5]

$$\frac{1}{2L} \sum_{i,j=1}^{N} \frac{k_i k_j}{2L} \delta_{c_i,c_j} = \sum_{c=1}^{n_c} \frac{1}{(2L)^2} \sum_{i,j\in C_c} k_i k_j = \sum_{c=1}^{n_c} \frac{k_c^2}{4L^2} , \tag{A.5}$$

where k_c is the total degree of the nodes in community C_c. Combining Eq. (A.4) and Eq. (A.5) gives the final result [5]

$$Q = \sum_{c=1}^{n_c} \left[\frac{L_c}{L} - \left(\frac{k_c}{2L} \right)^2 \right] .$$

(A.6)

A.1.2   Modularity change after merging two communities

Consider two communities A and B and denote with k_A and k_B the total degree of communities A and B respectively [5]. The change in modularity after merging the two communities is calculated as follows (using Eq. (A.6)) [5]

$$\Delta Q_{AB} = \left[\frac{L_{AB}}{L} - \left(\frac{k_{AB}}{2L} \right)^2 \right] - \left[\frac{L_A}{L} - \left(\frac{k_A}{2L} \right)^2 + \frac{L_B}{2L} - \left(\frac{k_B}{2L} \right)^2 \right] ,$$

(A.7)

where $L_{AB} = L_A + L_B + l_{AB}$ is the total number of edges in the merged community; L_A and L_B are the number of edges in communities A and B respectively and l_{AB} is the number of direct edges between communities A and B [5]. $k_{AB} = k_A + k_B$ is the total degree of nodes in the merged community [5]. Inserting the formulas for L_{AB} and k_{AB} in Eq. (A.7) gives

$$\Delta Q_{AB} = \frac{l_{AB}}{L} - \frac{k_A k_B}{2L^2} .$$

(A.8)

Appendix B

Network models: Derivations

B.1 The Erdős-Rényi model

B.1.1 Degree distribution

The probability that a node is connected to k other nodes and not to the $N - 1 - k$ others is given by

$$p^k(1 - p)^{N-1-k} .$$ (B.1)

The number of ways to choose the k nodes among the $N - 1$ possible candidates is given by the binomial coefficient

$$\binom{N - 1}{k} = \frac{(N - 1)!}{k!(N - 1 - k)!} ,$$ (B.2)

and the probability of being connected to exactly k other nodes becomes

$$p_k = \binom{N - 1}{k} p^k(1 - p)^{N-1-k} .$$ (B.3)

We thus derived the degree distribution of a random network.

B.1.2 Average degree

The expected mean degree is given by [52]

$$
\begin{aligned}
\langle k \rangle &= \sum_{k=0}^{N-1} k \, p_k \\
&= \sum_{k=0}^{N-1} k \binom{N - 1}{k} p^k(1 - p)^{N-1-k} .
\end{aligned}
$$ (B.4)

In order to simplify this equation we can make use of the following formula [52]

$$(p + q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} .$$ (B.5)

Differentiating both sides of this equation (with respect to p) gives us [52]

$$
\begin{aligned}
n(p+q)^{n-1} &= \sum_{k=0}^{n} \binom{n}{k} k p^{k-1} q^{n-k} \\
&= \frac{1}{p} \sum_{k=0}^{n} \binom{n}{k} k p^{k} q^{n-k} .
\end{aligned}
\tag{B.6}
$$

Finally, by substituting $q = 1 - p$, we get [52]

$$
\begin{aligned}
n(p+(1-p))^{n-1} &= \frac{1}{p} \sum_{k=0}^{n} \binom{n}{k} k p^{k} (1-p)^{n-k} \\
np &= \sum_{k=0}^{n} \binom{n}{k} k p^{k} (1-p)^{n-k} ,
\end{aligned}
\tag{B.7}
$$

where the right hand side is equal to Equation (B.4) with $n = N-1$ and we thus obtain $\langle k \rangle = (N-1)p$.

B.1.3   Poisson form of the degree distribution

We begin with the exact binomial distribution (Eq.(2.30))

$$
p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} ,
\tag{B.8}
$$

which characterizes a random graph. In the case $k \ll N$, we can simplify the binomial in the following way [4]

$$
\binom{N-1}{k} = \frac{(N-1)(N-1-1)...(N-1-k+1)}{k!} \approx \frac{(N-1)^k}{k!} .
\tag{B.9}
$$

We also note the following equality [4]

$$
\ln[(1-p)^{N-1-k}] = (N-1-k) \ln\left(1 - \frac{\langle k \rangle}{N-1}\right) ,
\tag{B.10}
$$

where we made use of the fact that $\langle k \rangle = (N-1)p$. Using the series expansion [4]

$$
\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - ... , \qquad \forall\, |x| \leqslant 1 ,
\tag{B.11}
$$

we get [4]

$$
\ln[(1-p)^{N-1-k}] \approx (N-1-k) \frac{\langle k \rangle}{N-1} = -\langle k \rangle \left(1 - \frac{\langle k \rangle}{N-1}\right) \approx -\langle k \rangle ,
\tag{B.12}
$$

which is valid if $k \ll N$. We thus find [4]

$$
(1-p)^{N-1-k} = e^{-\langle k \rangle} .
\tag{B.13}
$$

Substituting Eq. (B.13) and Eq. (B.9) in Eq. (B.8) and using $\langle k \rangle = (N-1)p$, we finally obtain the Poisson form of the degree distribution [4]

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} = \frac{(N-1)^k}{k!} p^k \mathrm{e}^{-\langle k \rangle}$$
$$= \frac{(N-1)^k}{k!} \left(\frac{\langle k \rangle}{N-1} \right)^k \mathrm{e}^{-\langle k \rangle} \tag{B.14}$$
$$= \mathrm{e}^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \ .$$

B.1.4   Diameter

The expected number of nodes at a distance d from a starting node is [4]

$$N(d) = 1 + \langle k \rangle + \langle k \rangle^2 + ... + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} \ . \tag{B.15}$$

$N(d)$ cannot exceed the number of nodes N in the network. The maximum distance or diameter d_{max} of the network can thus be found by setting [4]

$$N(d_{\mathrm{max}}) \approx N \ . \tag{B.16}$$

If $\langle k \rangle \gg 1$, we can neglect the -1 terms in Eq. (B.15), which gives [4]

$$N \approx \langle k \rangle^{d_{\mathrm{max}}} \ , \tag{B.17}$$

and thus the diameter of a random network becomes [4]

$$d_{\mathrm{max}} \approx \frac{\ln N}{\ln \langle k \rangle} \ . \tag{B.18}$$

B.2   The Watts-Strogatz model

B.2.1   The regular ring lattice

Clustering coefficient

A triangle in the regular ring lattice requires two edge traversals in the same direction and one in the opposite direction [43]. In the regular ring lattice every node has a fixed degree K, this means that each node is connected to $K/2$ neighbors on its left and $K/2$ neighbors on its right. It is then clear that the final, backwards step can span at most $K/2$ nodes [43]; otherwise we would end up at a different node from which we started, which would definitely not form a triangle.
Hence, the number of triangles for a given node is given by choosing the two forward target nodes from the $K/2$ possibilities [43]

$$\binom{\frac{K}{2}}{2} = \frac{1}{4} K \left(\frac{K}{2} - 1 \right) \ . \tag{B.19}$$

The number of connected triples per node is given by choosing two nodes out of the K neighbors [43]

$$\binom{K}{2} = \frac{1}{2}K(K-1) \, . \tag{B.20}$$

Since every node in the regular ring lattice is identical, we can easily find the total number of triangles and connected triples in the network by multiplying Eq. (B.19) and Eq. (B.20) with the number of nodes N. We then obtain the final result for the clustering coefficient in the regular ring lattice [43]

$$\langle C \rangle = \frac{3 \cdot N\frac{1}{4}K\left(\frac{K}{2}-1\right)}{N\frac{1}{2}K(K-1)} = \frac{3(K-2)}{4(K-1)} \, . \tag{B.21}$$

Average distance

The general formula for the average distance is given by Eq. (2.11)

$$l = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \, , \tag{B.22}$$

where the sum runs over the shortest path between every pair of nodes. Assume we have a regular ring lattice with N nodes and each node has K neighbors. Now we can calculate the distance between node i and every other node in the ring lattice

- There are K nodes adjacent to node i, so there are K nodes with a distance $d_{ij} = 1$

- There are K nodes that are neighbors of the neighbors of node i, so there are K nodes with a distance $d_{ij} = 2$

- ...

- There are K nodes that have a distance $d_{ij} = \frac{N}{K}$ between them and node i; its easy to see that these are the nodes that are the furthest away from node i

So the average distance for node i becomes [1]

$$l_i = \frac{1}{N-1}K\left(1+2+3+...+\frac{N}{K}\right) \approx \frac{K}{N}\left(1+2+3+...+\frac{N}{K}\right) \, , \tag{B.23}$$

where the last approximation holds for $N \gg 1$. Since every node in the regular ring lattice is identical we can easily find the average distance of the network [1]

$$l = \frac{1}{N} \sum_i l_i = \frac{1}{N}Nl_i \approx \frac{K}{N}\left(1+2+3+...+\frac{N}{K}\right) \, . \tag{B.24}$$

Using [1]

$$1+2+3+...+x = \sum_{i=1}^{x} i = \frac{x(x+1)}{2} \approx \frac{x^2}{2} \, , \tag{B.25}$$

where the last approximation again holds for $x \gg 1$; we finally obtain [1]

$$l \approx \frac{K}{N} \frac{\frac{N^2}{K^2}}{2} = \frac{N}{2K} \; . \tag{B.26}$$

Bibliography

[1] Chapter 9: Graphs and networks.

[2] E. Abbe. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, (18):1–86, Oct. 2018.

[3] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, Jan. 2002.

[4] A.-L. Barabási. Network science: Random networks. Sept. 2014.

[5] A.-L. Barabási. *Network Science*, chapter 9, pages 320–378. Cambridge University Press, July 2016.

[6] A. Barrat and M. Weigt. On the properties of small-world network models. *Europ. Phys. J. B*, 547(13), 2000.

[7] I. Benjamini, S.-O. Chan, R. O'Donnell, O. Tamuz, and L.-Y. Tan. Convergence, unanimity and disagreement in majority dynamics on unimodular graphs and random graphs. *Stochastic Processes and their Applications, Elsevier BV*, 126(9):2719–2733, Sept. 2016.

[8] E. Bozdag. Bias in algorithmic filtering and personalization. *Ethics and Information Technology, Springer Science and Business Media LLC*, 15(3):209–227, June 2013.

[9] M. Buchanan. *The Social Atom: Why the Rich Get Richer, Cheaters Get Caught, and Your Neighbor Usually Looks Like You*. BLOOMSBURY, 2007.

[10] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics, American Physical Society (APS)*, 81(2):591–646, May 2009.

[11] A. Clauset. Inference, models and simulation for complex systems: Lecture 13. Oct. 2011.

[12] A. Clauset. Network analysis and modeling, csci 5352: Lecture 6. 2017.

[13] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda. Complex systems: Analysis and models of real-world networks.

[14] L. da F. Costa. What is a complex network? (cdt-2). 2018.

[15] L. da F. Costa, O. N. O. Jr., G. Travieso, F. A. Rodrigues, P. R. V. Boas, L. Antiqueira, M. P. Viana, and L. E. C. da Rocha. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. 2007.

[16] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, Jan. 2007.

[17] A. S. da Mata. Complex networks: a mini-review. *Brazilian Journal of Physics*, 50(5):658–672, July 2020.

[18] G. Deffuant. Comparing extremism propagation patterns in continuous opinion models. *Journal of Artificial Societies and Social Simulation*, 9(3):8, 2006.

[19] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, June 2002.

[20] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World.* Cambridge University Press, 2010.

[21] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. S. Miguel, and V. M. Eguíluz. Is the voter model a model for voters? *Physical Review Letters*, 112(15), Apr. 2014.

[22] E. Ferrara. A large-scale community structure analysis in facebook. *EPJ Data Science*, 1(1), Nov. 2012.

[23] N. E. Friedkin and E. C. Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–205, 1990.

[24] N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. *Advances in Group Processes*, 16(1), Jan. 1999.

[25] T. Funke and T. Becker. Stochastic block models: A comparison of variants and inference methods. *PLOS ONE*, 14(4), Apr. 2019.

[26] W. Goddard and O. R. Oellermann. Distance in graphs. In *Structural Analysis of Complex Networks*, pages 49–72. Birkhäuser Boston, Sept. 2010.

[27] D. Gribel, T. Vidal, and M. Gendreau. Assortative-constrained stochastic block models. Apr. 2020.

[28] H. Heer, L. Streib, R. B. Schäfer, and S. Ruzika. Maximising the clustering coefficient of networks and the effects on habitat network robustness. *PLOS ONE*, 15(10), Oct. 2020.

[29] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[30] P. Holme and J. Saramäki. Temporal networks. *Physics Reports*, 519(3):97–125, Oct. 2012.

[31] H.-B. Hu and X.-F. Wang. Disassortative mixing in online social networks. *EPL (Europhysics Letters)*, 86(1), Apr. 2009.

[32] A. Ioannis and T. Eleni. Statistical analysis of weighted networks. 2007.

[33] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang. Attack robustness and centrality of complex networks. *PLOS ONE*, 8(4), Apr. 2013.

[34] B. Karrer and M. E. J. Newman. Stochastic blockmodels and community structure in networks. *Physical Review E, American Physical Society (APS)*, 83(1), Jan. 2011.

[35] M. Latapy and C. Magnien. Measuring fundamental properties of real-world complex networks.

[36] C. Lee and D. J. Wilkinson. A review of stochastic block models and extensions for graph clustering. *Applied Network Science, Springer Science and Business Media LLC*, 4(1), Dec. 2019.

[37] Y. Li, Y. Shang, and Y. Yang. Clustering coefficients of large networks. *Information Sciences*, 382-383:350–358, Mar. 2017.

[38] M. McGlohon, L. Akoglu, and C. Faloutsos. Statistical properties of social networks. In *Social Network Data Analytics*, pages 17–42. Springer US, 2011.

[39] L. Muchnik, S. Pei, L. C. Parra, S. D. S. Reis, J. S. A. Jr, S. Havlin, and H. A. Makse. Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific Reports*, 3(1), May 2013.

[40] M. E. J. Newman. Models of the small world: A review. *J. Stat. Phys.*, 101:819–841, 2000.

[41] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351, Sept. 2005.

[42] V. X. Nguyen, G. Xiao, X.-J. Xu, Q. Wu, and C.-Y. Xia. Dynamics of opinion formation under majority rules on complex social networks. *Scientific Reports, Springer Science and Business Media LLC*, 10(1), Jan. 2020.

[43] K. Pelechrinis. Telcom2125: Network science and analysis, 2015.

[44] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani. Activity driven modeling of time varying networks. *Scientific Reports*, 2(1), June 2012.

[45] N. Perra and L. E. C. Rocha. Modelling opinion dynamics in the age of algorithmic personalisation. *Scientific Reports*, 9(1), May 2019.

[46] B. Saha, A. Mandal, S. B. Tripathy, and D. Mukherjee. Complex networks, communities and clustering: A survey. 2015.

[47] A. Sîrbu, V. Loreto, V. D. P. Servedio, and F. Tria. Opinion dynamics: models, extensions and external effects. 2016.

[48] G. Thedchanamoorthy, M. Piraveenan, D. Kasthuriratna, and U. Senanayake. Node assortativity in complex networks: An alternative approach. *Procedia Computer Science*, 29:2449–2461, 2014.

[49] M. A. Turner and P. E. Smaldino. Stubborn extremism as a potential pathway to group polarization.

[50] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, June 1998.

[51] D. Zhang and G. Guo. A comparison of online social networks and real-life social networks: A study of sina microblogging. *Mathematical Problems in Engineering*, pages 1–6, 2014.

[52] J. Zhang, D. Goldburt, and J. Hopcroft. Lecture 01: Large graphs. Jan. 2006.